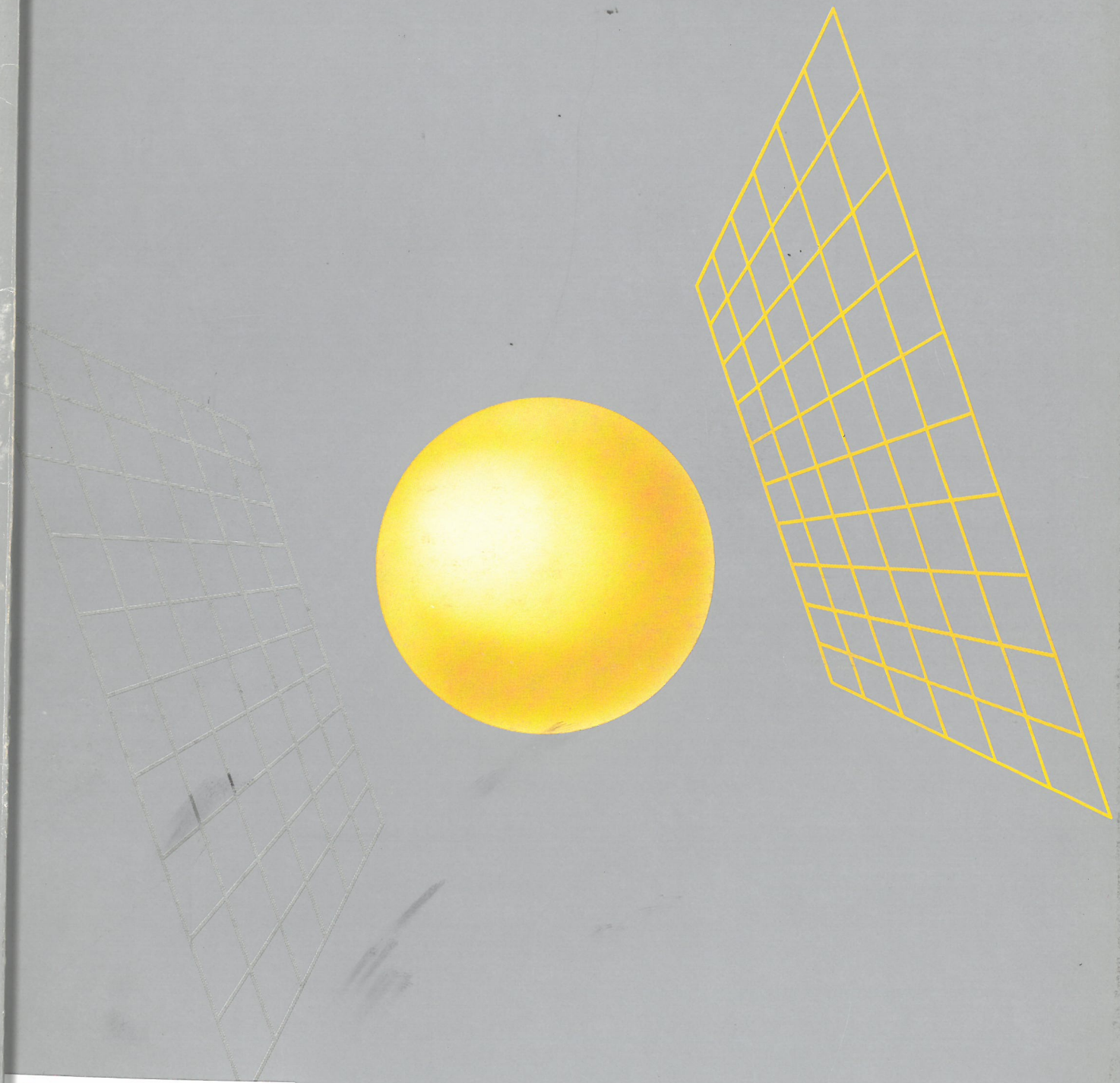
Quaderni di informatica 27

ssegna di tecnologia ed applicazioni degli elaboratori elettronici – Anno XII – Numero 2 – 1986



ARCHIVIO STORICO
ASSOCIAZIONE POZZO DI MIELE
FPDM-72
RQDI 125

CENTRO STUDI INFORMATICA E AUTOMAZIONE
Honeywell
Honeywell Information Systems Italia

Honeywell
Honeywell Information Systems Italia

Quaderni di informatica

Rassegna di tecnologia ed applicazioni degli elaboratori elettronici

Anno XIII - Numero 2 - 1986

Sommario

Orientamenti del mercato dell'informatica

L'articolo presenta una analisi sintetica e alcune fondamentali linee di tendenza del mercato dell'informatica.

C. PERETTI

L'automazione d'ufficio giudicata dai suoi protagonisti

L'automazione d'ufficio è un fenomeno che si sta dilatando a macchia d'olio, e che coinvolge, in varie forme, tutti i livelli aziendali.

G. OCCHINI

Intelligenza artificiale e discipline umanistiche

Le discipline umanistiche hanno vari punti di contatto con l'intelligenza artificiale e possono utilizzarne i concetti e le tecniche.

E. RICH

Tecnologie informatiche per i disabili

L'informatica può fare molto per i portatori di handicap. L'articolo illustra alcune soluzioni che consentono a persone con gravi limitazioni fisiche di comunicare e di svolgere un lavoro.

T. HILL, C. HUNT, R. YEATER

Dati sperimentali sullo sviluppo di grandi sistemi di software

Vengono riassunte le considerazioni più significative derivate da esperienze di sviluppo di grandi progetti di software.

A. PIZZARELLO

Una metodologia di misura della qualità del software

La qualità di un prodotto software può essere valutata dal produttore, prima che verificata dall'utente.

M. KARLIN

L'archeologia classica e l'elaborazione elettronica dei dati

L'archeologia sembrerebbe un tema molto lontano dal computer. L'Autore mostra, invece, come l'informatica possa essere utilmente impiegata anche in questa affascinante disciplina.

R. ROSATI

Direttore Responsabile
F. Filippazzi

Comitato di Redazione
A. Cicu, C. Falcetti, P. Lupo,
M. Nobile, G. Occhini
F. Sala

Redazione
Honeywell Information Systems Italia
Centro Studi Informatica e Automazione
Via G.M. Vida, 11 - 20127 Milano

Stampa
tipolito Maggioni s.a.s.

Autorizzazione del Tribunale di Milano
n. 93 del 20 Marzo 1974

«Quaderni di Informatica» ospita articoli e contributi di varia provenienza. La responsabilità delle opinioni espresse rimane agli Autori.

La riproduzione di articoli della rivista, o di parte di essi, è consentita solo citando la fonte e l'autore.

Orientamenti del mercato dell'informatica

CARLO PERETTI

Presidente Honeywell Italia
Milano

Il 1986 si apre, nel settore dell'informatica, all'insegna di una apparente contraddizione.

Da una parte si è parlato, per la prima volta nella storia del settore, di "crisi del computer" con riferimento alla flessione registrata nella fascia dei prodotti "personal".

Dall'altra, si assiste ad un crescente interesse del mercato per la cosidetta microinformatica di cui i prodotti tipo "personal" fanno parte.

Per rendersi conto di come ambedue questi fenomeni possano presentarsi contemporaneamente, è necessario un esame più approfondito della situazione a livello mondiale e quindi europeo ed italiano.

Consideriamo innanzitutto alcune caratteristiche della evoluzione del settore nel suo complesso (hardware, software, servizi) a partire dal 1980.

La tabella di fig. 1 mostra come, retrospettivamente, il tasso di sviluppo nel periodo si sia mantenuto superiore o uguale al 20% annuo.

Anche nel 1985, nonostante la cosiddetta "crisi", questo valore è rimasto sostanzialmente invariato, dimostrando così che le eventuali perdite in un'area sono state controbilanciate da guadagni compensativi in altre.

La tabella suggerisce però altre due considerazioni.

La prima che, anche per il futuro, gli analisti del mercato non prevedono mutamenti di tendenza a livello globale; la seconda che l'Europa, che ha espresso, nel trascorso quinquennio, un tasso di crescita ridotto rispetto alla media (11.2%), dovrebbe, in questo secondo scorcio degli anni '80, recuperare parte del ritardo.

Premesso che valutazioni a livello di aggregazione così elevato devono essere considerate in termini relativi più che assoluti, il fenomeno si può spiegare come conseguenza degli effetti indotti dall'andamento del dollaro e della congiuntura economica.

L'apprezzamento che la moneta americana ha avuto, negli ultimi anni, su quella europea e l'insoddisfacente ripresa economica di certi paesi possono giustificare il tasso di crescita inferiore che si è registrato in Europa; l'inversione di tendenza che, per ambedue i fattori, sembra si stia oggi verificando, è all'origine delle previsioni di recupero.

(*) Questo articolo fa parte dei saggi raccolti nel volume *«Italia '86 – Rapporto di Primavera»*, edizioni Euritalia.

Fig. 1 - Il mercato mondiale dell'industria informatica negli anni '80 (Fonte: InfoCorp)

	1980		1983		1985		1987		1990	
	B$	%	B$	%	B$	%	B$	%	B$	%
USA	26.6	45.6	47.3	50.4	71.0	53.9	98.2	52.8	139.9	49.9
EUROPA	20.3	34.8	27.3	29.1	34.6	26.3	50.0	26.9	81.2	29.0
AREA PACIFICO	7.2	12.3	12.7	13.5	17.2	13.1	25.2	13.5	39.8	14.2
RESTO DEL MONDO	4.3	7.3	6.5	6.9	8.8	6.7	12.6	6.8	19.4	6.9
TOTALE	58.4	100.0	93.8	100.0	131.6	100.0	186.0	100.0	280.3	100.0

Piuttosto anomala si presenta, alla luce di queste considerazioni, la situazione italiana.

In Italia, sia nel 1984 che nel 1985, il tasso di crescita registrato nel settore informatico si aggira sul 25%, nettamente superiore quindi a quello medio europeo pur tenendo conto del rapporto lira/dollaro.

Ma il progresso si è registrato in tutte le fascie di prodotti e servizi per cui, tornando all'argomento iniziale, neanche per il personal (almeno quello di tipo professionale) il 1985 è stato un anno deludente.

Questa differenza di comportamento del nostro mercato rispetto a quello di altri paesi può spiegarsi con una analisi più approfondita delle modalità di utilizzo della apparecchiatura.

Precisamente, il personal ha oggi sostanzialmente due tipi di mercato tra loro molto differenti. Il primo è quello degli studi professionali, delle piccole imprese artigiane, dei commercianti ecc. dove viene utilizzato come elaboratore a basso costo per risolvere i tradizionali problemi di gestione.

In questo mercato, in sostanza, il personal ha segnato un semplice, anche se quantitativamente molto significativo, allargamento della base di utenza, senza modificare l'impostazione della informatica che potremmo definire "classica".

Il secondo mercato, potenzialmente ancora più interessante, ma, in termini di volumi, oggi più contenuto, è quello della informatica individuale nelle aziende dove già esiste una consolidata struttura EDP.

Il distinguere con chiarezza questi due tipi di domanda rappresenta la chiave interpretativa di ciò che sta avvenendo.

Nel rapporto edito annualmente dalla HISI sulla evoluzione del mercato EDP in Italia si ricorda come l'andamento nel tempo del tasso di penetrazione di una nuova tecnologia, segua sempre una curva evolutiva dal tipico profilo ad S.

Questa curva esprime il fatto che, all'inizio, il fenomeno procede con ritmi relativamente lenti per le difficoltà di acquisizione, successivamente segue una fase di rapida crescita che poi tende di nuovo a rallentare con l'avvicinarsi della saturazione.

Ebbene, la figura 2 mostra come nel campo della piccola utenza, il mercato italiano presenti un numero rilevante di settori nel tratto centrale della curva evolutiva che è appunto quella di più rapido sviluppo.

Quantitativamente, la crescita complessiva del parco informatico presso i piccoli utenti è stata in questi ultimi tempi superiore al 25% annuo in numero di macchine e al 18% in valore con punte che, nel caso degli autotrasportatori o delle agenzie di viaggio, sono state ancora più alte (attorno al 40%).

Questa constatazione giustifica il persistente "boom" del personal nel mercato della piccola utenza italiana ma anche il rallentamento verificatosi negli Stati Uniti.

In Usa, dove il personal è entrato prima e dove la cultura informatica è più diffusa, si è già arrivati, in questa fascia di mercato, al terzo tratto della curva, quello cioè più prossimo alla saturazione.

Per quanto riguarda il secondo tipo di utenti costituito dalle aziende di medie e grandi dimensioni, lo stesso rapporto HISI dimostra come la penetrazione del personal sia più rapida dove già esiste un sistema informatico ben strutturato, ma comunque comporti un radicale cambiamento nell'approccio tradizionale.

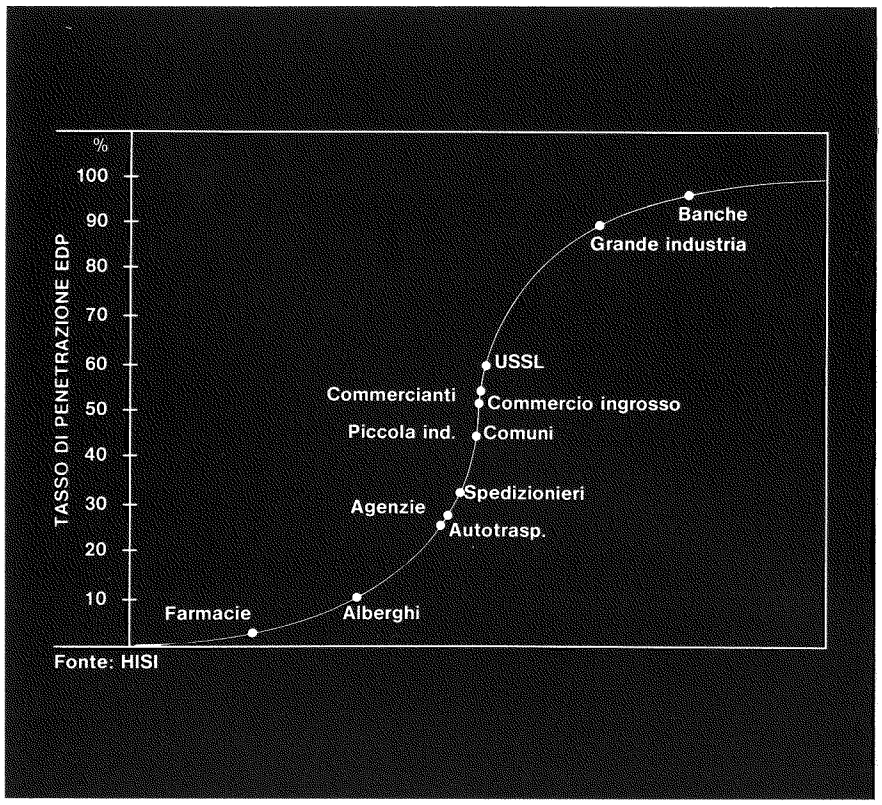


Fig. 2 - Tasso di penetrazione della microinformatica per settore (Fonte: Honeywell I.S.I.)

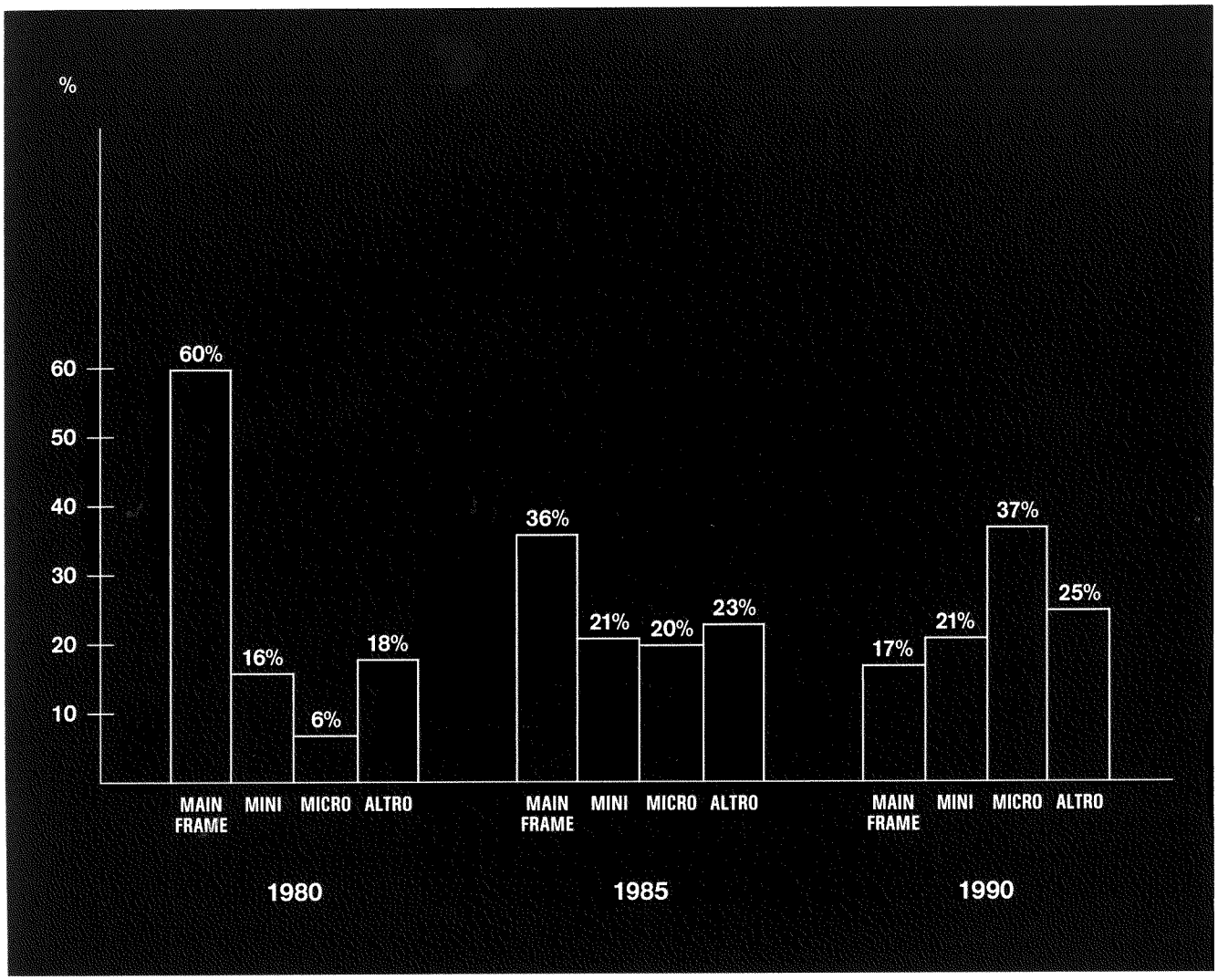


Fig. 3 - Percentuali di contribuzione al fatturato delle principali classi di prodotti informatici (Fonte: IDC).

Cambiamento che non è solo tecnico ma anche, per così dire, "culturale", in quanto presuppone che ruoli e strutture sistemistiche e organizzative consolidati vengano rimpiazzati da nuove realtà tuttora in via di definizione.

Più del 60% degli utenti di "mainframe" risulta oggi aver installato dei personal, ma in misura ancora estremamente limitata.

In altre parole, il personal nella azienda o, più generalmente, la microinformatica come tecnologia di sviluppo della produttività individuale, è ancora nella prima fase della curva evolutiva.

Il potenziale di mercato è, in questa area, ancora più rilevante di quello dei nuovi utenti, ma si svilupperà con gradualità perchè è condizionato dall'evolversi di complesse strutture e questo vale in Italia, come in Europa e negli Stati Uniti.

La fig. 3, in parte retrospettiva e in parte previsiva, dimostra appunto come nel decennio in corso, stiamo assistendo a un rovesciamento del tradizionale rapporto tra mainframe (grandi elaboratori), mini e microelaboratori in termini di fatturato.

Dalla situazione del 1980 in cui i grandi elaboratori assorbivano una percentuale maggioritaria del mercato, ci si sta avviando a una situazione in cui contribuiranno in modo relativamente minoritario.

La crisi di cui si è tanto parlato nell'ultimo anno non è quindi altro che il riflesso di un fenomeno di riassestamento generale del settore in termini di fascie di prodotto e quindi di categorie di utilizzatori.

Questo riassestamento non comporta, ovviamente, una riduzione in termini di valore assoluto del mercato dei grandi sistemi, che è anzi in continua crescita, né tantomeno un ridimensionamento della loro funzione baricentrica.

In effetti, lo sviluppo accelerato della microinformatica, che è all'origine di questo riassestamento, si inquadra nel diffondersi della filosofia dei sistemi di elaborazione distribuiti in cui coesistono, in modo integrato, classi diverse di prodotti, ciascuna con un suo preciso ruolo e con sue precise tipologie di utilizzo.

L'automazione d'ufficio giudicata dai suoi protagonisti

GIULIO OCCHINI

Honeywell Information Systems Italia
Centro Studi Informatica e Automazione
Milano

1. Introduzione

A chi si rivolge la automazione ufficio e, più in generale, l'informatica individuale?

La risposta più banale potrebbe essere: a tutti coloro che, indipendentemente dalla organizzazione di cui fanno parte, devono, nel loro lavoro, trattare informazioni, nel senso di accumularle, selezionarle, elaborarle, presentarle.

Ma questo non basta per determinare le priorità con cui effettuare gli investimenti e decidere della loro entità.

A questo proposito anzi non ci si deve fare ingannare dalla cifra relativamente modesta necessaria per l'automazione del singolo posto di lavoro; se è vero che l'informatica, come qualsiasi altra tecnologia, diventa redditizia al di sopra di una certa soglia critica di utilizzo, l'investimento richiesto a una azienda per la sua diffusione a livello individuale assume rapidamente dimensioni tali da richiedere una accurata pianificazione e una approfondita valutazione del rapporto costo/beneficio.

È necessario quindi una analisi più puntuale del lavoro di trattamento informazioni.

Qualunque sia l'organizzazione, lo schema di Fig. 1 distingue, nell'ambito di chi esercita una tale attività (*), coloro che svolgono compiti di tipo esecutivo da quelli che, a livello specialistico e manageriale, esercitano una più o meno ampia capacità discrezionale (i cosiddetti "knowledge worker").

Nell'ambito della prima categoria è possibile ulteriormente distinguere tra gli "addetti alle transazioni", vale a dire coloro che elaborano informazioni secondo procedure formali, e "personale di supporto" che ha il compito di rendere più produttivo il lavoro dei "knowledge worker" e che comprende le funzioni di segreteria, copisteria, etc.

Le procedure formali secondo cui operano gli addetti alle transazioni, sono per definizione, programmabili, ed è alla loro automazione che storicamente si è rivolto l'EDP aziendale con il proposito di sostituirsi all'uomo e quindi di affrontare lo stesso lavoro con minori risorse o, più frequentemente, moli di lavoro maggiore a parità di organico.

L'esperienza degli Istituti di credito che, per primi, in modo sistematico, hanno percepito e valorizzato le potenzialità di un tale impiego della tecnologia informatica, è, da questo punto di vista, esemplare; ma risultati analoghi si sono ottenuti anche nelle aziende di tipo industriale e commerciale.

Il sistema EDP aziendale, che tende a coprire man mano le varie funzioni di tipo strutturato che si svolgono nella azienda, impone, per il fatto stesso di costituirsi, una logica di razionalizzazione organizzativa cui devono conformarsi gli utenti pena l'insuccesso del sistema stesso; di qui il nome di "informatica di organizzazione" che talvolta si assegna a questo processo di automazione.

Questo tipo di informatica non è però applicabile, proprio per le sue caratteristiche, al lavoro dei "knowledge worker" e tanto meno quanto più alto è il livello di discrezionalità da essi esercitato.

Di qui la necessità di procedere, nei riguardi di questi ruoli, con un approccio sostanzialmente diverso.

L'informatica individuale che si avvale della moderna stazione di lavoro di ufficio (solo in parte riconducibile al personal) ha proprio come obiettivo di estendere la possibilità di impiego della tecnologia ai ruoli direttivi e professionali della azienda; l'obiettivo è sempre quello di migliorarne la prestazione ma questa volta in termini soprattutto qualitativi ("meglio" piuttosto che "di più", "efficacia" rispetto ad "efficienza").

2. Importanza dei knowledge worker

Per apprezzare l'importanza del ruolo esercitato dai "knowledge worker" nelle moderne organizzazioni, torna utile partire dallo schema di Fig. 2.

Essa fornisce una indicazione di larga massima su come oggi si ripartisca il costo del lavoro di ufficio in una azienda manufatturiera che operi in un settore tecnologicamente avanzato (cioè con notevole incidenza delle componenti di progettazione e di marketing) oppure in una società di servizi (tipicamente finanziari).

(*) Le categorie schematizzate in figura non sono sempre di agevole individuazione come dimostra lo studio di M. Porat sul settore da lui battezzato "dell'informazione".

Fig. 1 - Il lavoro di trattamento informazioni e le due informatiche.

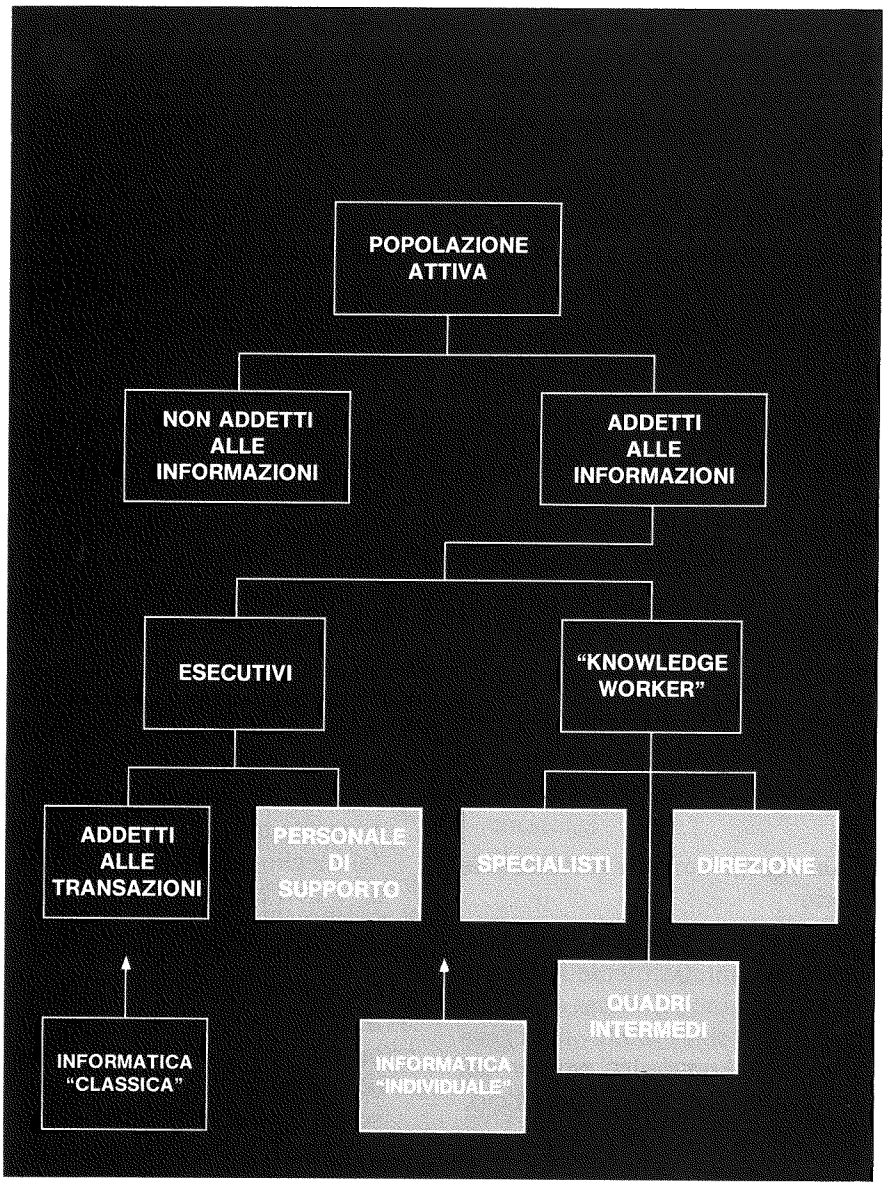
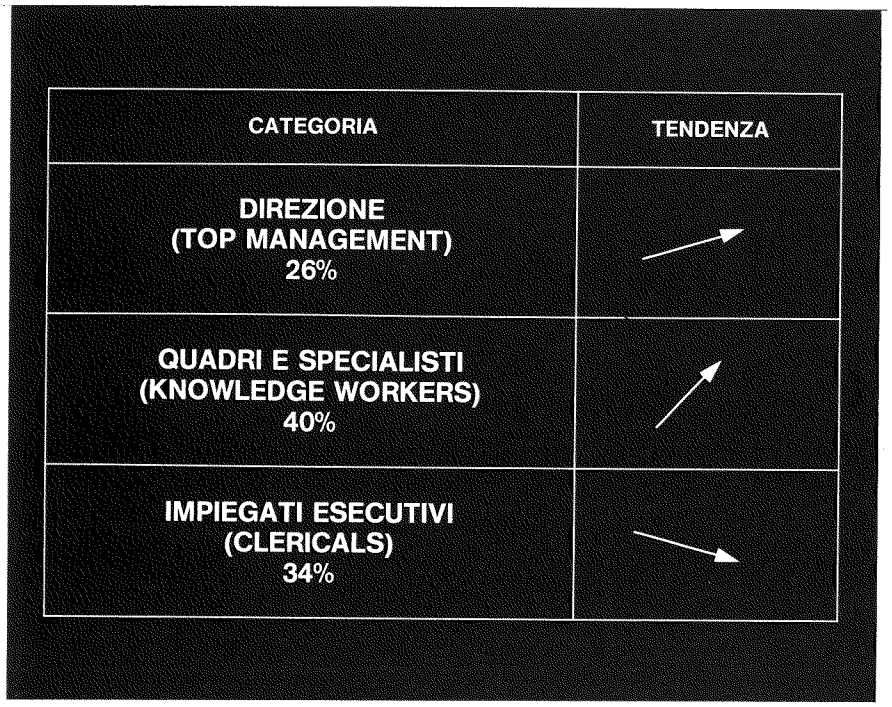


Fig. 2 - Ripartizione e tendenze del costo del lavoro di ufficio.



La ripartizione mostra che, al di là del valore delle singole percentuali, variabili da azienda ad azienda, la quota di lavoro esecutivo è oggi minoritaria e tendenzialmente in riduzione.

Solo 15-20 anni fa la stessa rilevazione avrebbe infatti mostrato come il lavoro esecutivo fosse ancora preddominante a scapito di quello professionale e direttivo.

Il cambiamento che sta avvenendo ha rilevantissime conseguenze di carattere sociale ed economico ed è stimolato da una tecnologia, come quella informatica che, prima, ha consentito di "svuotare", automatizzandoli, i compiti esecutivi, favorendo così lo sviluppo, nell'ambito di una globale compatibilità di spesa, delle attività con contenuto discrezionale e, adesso, sta appunto orientandosi a queste ultime per migliorarne l'efficacia.

3. Knowledge worker e automazione ufficio

Nel prossimo paragrafo verranno presentati sinteticamente i risultati di una indagine promossa dalla Honeywell e riguardante l'atteggiamento dei knowledge worker nei confronti dell'automazione ufficio.

Ma prima di passare alle indicazioni statistiche è forse opportuno anticipare alcune considerazioni d'inquadramento che danno conto del ruolo traente che questa categoria professionale sta esercitando nel processo di innovazione.

a) Il settore amministrativo della azienda ha storicamente giocato un ruolo cruciale nell'indirizzare l'informatica di organizzazione.
   In una situazione economica ragionevolmente stabile, d'altra parte, la funzione amministrativa è giustamente dominante dato che l'obiettivo è di ridurre i costi operativi a parità di prodotto/servizio fornito.
   Ma se, come accade oggi, si viviono periodi di accelerate trasformazioni e quindi anche di acuta crisi, la componente "ingegneristica" della azienda (che, si badi bene, non è assolutamente limitata al settore Ricerca e Sviluppo ma si estende a tutte le aree funzionali ed è caratterizzabile come capacità creativa e disponibilità al nuovo), tende ineluttabilmente a prevalere come carta vincente per compiere, attraverso la innovazione, veri e propri salti di qualità nella efficacia operativa della azienda.
   Il baricentro di potere tende di conseguenza a spostarsi dall'area puramente gestionale a quella progettuale e questo contribuisce a far si che il taglio della informatica di organizzazione si dimostri sempre meno adeguato sia come capacità di risolvere i problemi che come tempi di risposta.

b) Il decentramento operativo con i suoi strumenti di autonomia, di delega, etc. è da sempre la filosofia organizzativa che tende a prevalere in situazioni di alta competitività, in quanto è l'unica che consenta prestazioni ottimizzate "al meglio delle possibilità locali" e non più "nella media".
   Un tale decentramento non va visto come necessariamente antinomico al controllo e alla centralizzazione delle decisioni strategiche, specie oggi che le reti telematiche consentono di prescindere dalla localizzazione dei dati, ma è tuttavia uno stimolo ad assumere iniziative locali in modo indipendente e in base a finalità locali.

c) Sta ormai penetrando nella cultura economica il concetto che la risorsa "informazione" sia la più conveniente per gestire, in situazioni di incertezza, una qualsiasi organizzazione complessa.
   L'alternativa alla informazione è infatti rappresentata dall'accantonamento di riserve (materiale, personale, denaro) mobilitabili al sopraggiungere dell'imprevisto (e imprevedibile, dato il generale contesto di incertezze) per poterlo fronteggiare e superare.
   Ma questa strategia, dati gli attuali costi di immobilizzo, è in pratica perdente dal punto di vista economico rispetto a quella di mettere in atto un sistema informativo (e decisionale) sufficientemente flessibile che si adatti dinamicamente al cambiamento e, anzi, metta in grado chi decide di trarre da esso il massimo di opportunità.
   L'informatica di organizzazione, che dovrebbe riguardare i fatti relativamente "invarianti" della gestione, si presenta intrinsecamente inadatta a far fronte a una esigenza di elevata reattività (questa, della mancanza di flessibilità, è in effetti una delle critiche più frequenti e motivate alla tradizionale struttura dell'EDP aziendale); l'informatica individuale dovrebbe proprio essere finalizzata a questo scopo anche perchè è assai più vicina (con capacità di condizionamento) allo sviluppo del processo decisionale.

4. L'indagine

Finalità e metodologia

In quanto creano, archiviano, ricercano ed elaborano informazioni per decidere, i knowledge worker vengono ad essere profondamente coinvolti nel processo di "informatizzazione" del lavoro e, in particolare, da quella sinergia tra tecnologie informatiche e di comunicazione che va sotto il nome di automazione ufficio.

L'atteggiamento con cui questa categoria professionale guarda alle nuove tecniche è quindi cruciale per la loro diffusione ed è appunto per appurarlo che la Honeywell ha promosso una ricerca presso le principali Corporation statunitensi (*).

Precisamente, sono state definite come tali le organizzazioni industriali, di servizio o finanziarie che figurino tra le prime 500 (secondo la classifica pubblicata in FORBES, April 30, 1984) in una o più delle seguenti categorie:

- Capitale
- Valore di mercato
- Profitti
- Vendite

(*) *Office automation and the workplace: a national survey of knowledge workers. Honeywell Technalysis, november 1984.*

L'indagine ha riguardato 701 knowledge worker distribuiti uniformemente nelle 7 aree funzionali seguenti:

- Ricerche, studi e progettazione
- Finanza
- Affari Legali
- Personale
- Produzione
- Marketing
- Acquisti

Una volta scelte, con criteri rigorosamente statistici, le persone da intervistare nelle varie Corporation, ciascun intervista ha richiesto circa 20 minuti di colloquio individuale.

a) Argomenti generali

La maggior parte degli intervistati dispone già di strumenti informatici per il suo lavoro quotidiano:

- 53% di personal (tipo professionale)
- 65% di terminale
- 75% di word processor

Dalle risposte è risultata una correlazione molto stretta tra aumento di produttività e impiego del computer.

Prima ancora che si rendessero conto degli scopi dell'indagine, agli intervistati veniva richiesto come avrebbero utilizzato 10.000 $ se l'azienda li avesse messi loro a disposizione per aumentare la produttività.

Un sorprendente 83% ha dichiarato che li avrebbe spesi in strumenti o tecniche informatiche e solo una piccola minoranza ha fatto riferimento ad aumenti di organico o ad altri aspetti quali la formazione.

Significativo è anche il fatto che con "automazione ufficio" la maggior parte degli intervistati intenda la diffusione di strumenti informatici; altre tecniche come la telefonia evoluta o la telecopiatura sono considerate del tutto marginali.

b) Atteggiamento nei riguardi della automazione ufficio

La gran parte degli intervistati ha espresso un notevole apprezzamento nei riguardi della automazione ufficio, che in generale coincide per loro con l'informatica.

Precisamente:

- l'81% dichiara di avere personalmente tratto vantaggi dall'uso di strumenti di automazione ufficio.
- il 93% pensa di essere in grado di formulare decisioni migliori e più argomentate
- il 94% afferma di disporre di più tempo per attività creative come la pianificazione o la valutazione di alternative di condotta
- il 92% è d'accordo sull'affermazione che l'automazione ufficio crea opportunità per chi disponga di buone capacità professionali.

Specialmente nel caso di impiego di personal e di word processor, gli interrogati ritengono di averne tratto significativi vantaggi.

Anche il 71% di quelli che, al momento, non hanno a disposizione strumenti informatici, esprime vivo interesse ad utilizzarli in futuro.

c) Pervasività della automazione

Word processor, personal e terminali sono ormai diventati elementi standard della dotazione di scrivania del knowledge worker.

Tre su quattro dispongono di word processor e il 68% di terminali; minore, ma tuttavia significativa (53%), anche la presenza di personal.

Il personal lo si trova con maggior frequenza nei settori finanziari e di marketing, il terminale nell'area di ricerca e di progettazione e il word processor negli uffici legali.

La maggior parte degli intervistati (dal 50% al 62%) dichiara di utilizzare lo strumento informatico "frequentemente" e di trarre vantaggio dal suo impiego. Meno del 20% lo utilizza "raramente" o "mai".

Circa la metà (51%) lo usa da più di due anni e tra questi, più del 70% non ha avuto difficoltà nell'apprenderne le modalità operative.

Sebbene siano assai meno diffuse degli elaboratori o dei terminali, anche le nuove tecnologie di comunicazione sono presenti negli uffici automatizzati.

Il 53% dei knowledge worker può accedere a centri di teleconferenza (ma solo il 30% ne fa uso) e il 45% si serve di "telefoni intelligenti" con la capacità di comporre automaticamente i numeri più frequentemente chiamati.

Una minoranza significativa (41%) si avvale di servizi di posta elettronica e di scambio di messaggi.

d) Valutazione delle singole tecnologie

I knowledge worker, sebbene favorevoli a tutte le tecnologie di automazione ufficio, lo sono in modo differenziato con una netta preferenza per quelle facenti capo all'informatica.

Precisamente il giudizio di merito risulta il seguente:

- Word processing	77%
- Personal computing	71%
- Terminale	69%
- Telefono programmabile	45%
- Posta elettronica	41%
- Centro di smistamento messaggi	37%
- Teleconferenza	30%

e) Vantaggi riscontrati

Gli intervistati presentano una notevole concordanza di vedute sui vantaggi della automazione.

Le poche differenze riscontrate riflettono infatti le esigenze diverse dei vari tipi di lavoro.

Ad esempio, gli addetti al servizio legale usano più il word processor mentre, come già detto, nel marketing e nella direzione del personale, lo strumento più frequentemente utilizzato è il personal.

Pressochè tutti sono comunque convinti che il vantaggio principale stia nella riduzione dei tempi operativi: l'informatica, cioè, consentirebbe di far più lavoro nello stesso tempo.

Quasi tutti (85%) pensano inoltre (ma non sono in grado di provarlo in modo altrettanto sicuro) che l'informatica consenta loro di fare un lavoro migliore anche in termini qualitativi.

f) Software

All'entusiasmo per i vantaggi derivanti dall'impiego delle tecnologie di automazione ufficio, non si ac-

compagna un atteggiamento altrettanto positivo nei riguardi del software disponibile sul mercato.

La maggior parte (>80%) ritiene infatti che il software sia solo qualche volta in grado di corrispondere adeguatamente alle necessità.

Nella più parte dei casi, essi stessi in prima persona e le loro aziende sono stati costretti a sviluppare programmi "ad hoc" per sostituire o, quanto meno, personalizzare l'offerta di mercato esistente.

g) Utilizzatori

Alcuni trovano l'automazione ufficio più utile di altri per il loro lavoro.

Si stabilisce, in questo, una diretta correlazione tra tempo dedicato all'uso della apparecchiatura e grado di apprezzamento positivo della sua utilità.

Il profilo dell'utilizzatore più entusiasta suggerisce che si tratti di:

- una persona relativamente giovane
- che si avvale dell'automazione ufficio per un tempo che varia tra 2 e 4 ore giornaliere
- che se ne avvale da due anni o più
- che opera nel settore finanziario, in quello tecnico o nella produzione.

Non si registrano differenze significative tra uomini e donne.

h) Modalità di introduzione

Il modo con cui l'azienda ha avviato l'automazione ufficio influisce evidentemente sull'atteggiamento dei dipendenti.

In generale, i knowledge worker danno un giudizio positivo sulla strategia di introduzione adottata dall'azienda; 7 su 10 pensano infatti che la direzione si sia posta obiettivi realistici e che l'azienda tenga conto delle esigenze dei dipendenti.

Molti (75%) ritengono che l'azienda abbia attentamente valutato le opinioni di persone come loro nella scelta delle apparecchiature.

Tra le varie aree aziendali, quella che manifesta l'atteggiamento più critico è il marketing ed è nel marketing che si rileva il maggior scetticismo relativo nei riguardi della automazione.

Coloro che lavorano nel marketing esprimono un atteggiamento ambivalente su tutto il proceso di informatizzazione e sul modo con cui viene gestito aziendalmente.

Essi si avvalgono degli strumenti più raramente degli altri e per minor tempo e sono quelli più inclini a ritenere che la loro azienda:

- abbia opposto resistenza all'uso dell'informatica (15% contro meno della metà negli altri settori)
- si preoccupi più di comprare macchine che di coloro le dovranno usare
- nutra aspettative irrealistiche su ciò che l'automazione può e non può fare.

i) Non utilizzatori

Uno su quattro dei knowledge worker che non hanno accesso a tecnologie informatiche si dimostra altrettanto entusiasta dei suoi colleghi sulle prospettive dell'automazione.

Sette su dieci dichiarano di essere "molto" o "in qualche modo" interessati ad utilizzare un elaboratore e i vantaggi che si ripromettono sono sostanzialmente gli stessi di quelli dichiarati dai loro colleghi: maggiore velocità di ciclo, reperimento più rapido delle informazioni e quindi aumento della produttività.

l) Modalità di impiego

Indipendentemente dalla mansione, i knowledge worker tendono a utilizzare l'automazione ufficio secondo modalità comuni.

La maggior parte (58%) passa meno di due ore per giorno al terminale (il 23% tra 2 e 4 ore) e lo fa per reperire informazioni da basi di dati locali o centrali (82%), per analizzare cifre (73%) o per comunicare con altri utenti (43%).

5. Conclusioni

Per quanto la realtà del mondo del lavoro statunitense presenti significative differenze rispetto a quella europea e particolarmente italiana, essa costituisce tuttavia, e non solo nel campo della automazione, un riferimento obbligato per il mondo occidentale.

In particolare, il ruolo portante economico e sociale che la categoria dei knowledge worker ha da tempo assunto negli Stati Uniti, se lo stanno conquistando da qualche anno, anche in Italia, i quadri e gli specialisti.

L'innovazione tecnologica, e nella fattispecie informatica, non è estranea a questo fenomeno; la disponibilità al nuovo dei ruoli tecnici e di management intermedio li qualifica naturalmente a prendere la guida dei processi di trasformazione e a provocare quello spostamento del baricentro di potere verso compiti progettuali cui si è accennato al paragrafo 3.

Da questo punto di vista l'indagine Honeywell non fa che confermare la forte motivazione dei knowledge worker statunitensi a interpretare con coerenza il ruolo di agenti di cambiamento che è stato a loro affidato.

Intelligenza artificiale e discipline umanistiche

ELAINE RICH

University of Texas
Austin (USA)

1. Introduzione

L'Intelligenza Artificiale (I.A.) si può definire l'applicazione dell'intelligenza dell'uomo al computer per renderlo altrettanto capace [5].

Essa comprende pertanto attività quali la soluzione di problemi, la comprensione del linguaggio naturale, la percezione del mondo fisico, le applicazioni scientifiche. Che cosa ha dunque a che fare l'Intelligenza Artificiale con le discipline umanistiche? Ci possono essere due livelli di risposta a questa domanda. Anzitutto, l'intelligenza artificiale si basa principalmente sulla elaborazione simbolica, ossia su simboli in generale anzichè su numeri. Poichè gli oggetti che vengono manipolati nello studio delle discipline umanistiche sono appunto simboli, ci possono essere delle tecniche di elaborazione che i due campi possono condividere.

Essi possono tuttavia compenetrarsi anche più profondamente. L'intelligenza artificiale cerca di risolvere problemi inerenti al mondo in cui viviamo. Molti di essi si incentrano sui popoli e sulle culture da essi create.

Probabilmente la classe più importante di questi problemi riguarda il linguaggio, ma ci sono altri problemi che esigono la comprensione del modo con cui gli esseri umani affrontano i problemi che sono chiamati a risolvere; tale comprensione può essere utilizzata per definire che tipo di conoscenza dovrebbe avere una macchina per risolvere gli stessi problemi.

Oltre a ciò, ogni problema richiede in pratica una comprensione della natura della conoscenza e dei meccanismi del ragionamento. Gli umanisti lavorano proprio a risolvere molti problemi di questo tipo; c'è pertanto spazio per una interazione profonda tra l'intelligenza artificiale e svariate discipline nell'ambito delle scienze umanistiche.

2. Tecniche di programmazione comuni

Dai punti di contatto tra I.A. e discipline umanistiche, si possono trarre diverse indicazioni sulla possibilità di usare tecniche di programmazione comuni. Le ricerche di I.A. hanno dato impulso allo sviluppo di linguaggi di programmazione ad alto livello, concepiti per l'elaborazione simbolica. Il LISP [11] è il linguaggio più importante, ed è stato in particolare utilizzato nello sviluppo dei sistemi interattivi.

Tali sistemi sono particolarmente utili in quelle applicazioni (ad esempio, analisi stilistica o progetto architettonico) che richiedono una stretta interazione tra l'uomo e il programma.

In tempi recenti, con l'avvento di stazioni di lavoro individuali che supportano il LISP, e di sistemi grafici di media precisione, entrambi integrati in un unico ambiente di programmazione, lo sviluppo di programmi simbolici interattivi è diventato ancora più agevole.

Se, prescindendo dalle caratteristiche generali del nostro programma, prendiamo in esame le tecniche specifiche usate, scopriamo un ulteriore terreno in comune. Si consideri anzitutto il "parsing", cioè il processo di analisi che si applica agli oggetti per determinarne le strutture. I linguisti tengono in gran conto il parsing come strumento per l'assegnazione di strutture alle frasi; ed i musicologi si affidano al parsing per assegnare strutture alle creazioni musicali. Poichè anche nella I.A. occorre strutturare la rappresentazione di oggetti del mondo reale, pure essa deve fare grande affidamento su tecniche di parsing, che svolgono un ruolo importante nella comprensione del linguaggio naturale e costituiscono la base di molti programmi per il riconoscimento delle forme ("pattern recognition"). Talvolta la descrizione strutturale completa che si otterrebbe analizzando un oggetto a fronte di una grammatica, risulta non necessaria o non fattibile. Quando ciò accade, si rivela spesso utile una forma meno potente di analisi, il "pattern matching" (accoppiamento di forme). Gli studiosi di letteratura utilizzano il matching nella loro ricerca su configurazioni specifiche di parole e di frasi. Nella I.A., il matching è importante ad esempio nell'ambito del riconoscimento della voce (dove il livello dei disturbi nel segnale preclude spesso l'utilizzo delle tecniche formali di parsing), come pure nel reperire informazioni utili sul mondo reale a partire dalla base di conoscenza disponibile.

Nonostante l'utilità di poter condividere determinate tecniche di programmazione (come ad esempio LISP o parsing), non è in questo campo che l'I.A. e le discipline umanistiche rivelano più compiutamente la loro affinità concettuale.

Il seguito dell'articolo illustrerà questo rapporto più profondo prendendo in esame un problema specifico di I.A. e mostrando come la sua soluzione dipenda dalla conoscenza degli esseri umani e del mondo in cui essi vivono.

3. Un problema di intelligenza artificiale

Si consideri il problema costituito dalla lettura del seguente paragrafo e dal dover dare risposta a domande ad esso relative:

«L'Azienda CN è preoccupata dello andamento dei profitti nel primo trimestre ed ha deciso di lanciare una campagna pubblicitaria a tappeto con lo scopo di promuovere la sua nuova linea di accumulatori e di candele per auto. Ha ingaggiato allo scopo la ditta locale L & C, che l'anno scorso condusse una campagna promozionale molto riuscita per conto di uno dei principali concorrenti della CN. Il presidente della CN, H.R. Smith è ottimista circa il successo dei suoi nuovi sforzi. Egli fonda questo ottimismo su statistiche recenti che mostrano come il numero di veicoli nel paese cresca a ritmo costante del 5% all'anno. L'annuncio della nuova campagna è giunta come una buona nuova per gli abitanti di Tube City, dove si trovano le principali unità di produzione della CN.»

Chi abbia letto questa notizia, potrebbe rispondere come segue a domande in merito:

1. «Per quale motivo la CN sta lanciando una nuova campagna pubblicitaria?»

   «Perchè vuole migliorare i profitti aumentando le vendite».
2. «Perchè Smith pensa che la sua iniziativa avrà successo?»

   «Perchè l'aumento del numero di veicoli significa che ci sarà anche una maggiore domanda di accumulatori e di candele».
3. «Perchè è insoddisfatto il management della CN?»

   «Perchè i profitti del primo trimestre sono stati inferiori alle aspettative e si teme che anche i profitti futuri saranno bassi».
4. «Perchè gli abitanti di Tube City sono contenti della nuova campagna pubblicitaria?»

   «Perchè temevano di perdere il posto qualora la CN riducesse la produzione».

Se vogliamo costruire un programma che possa "comprendere"(*) questa storia e rispondere a domande in merito, dobbiamo fornigli una adeguata base di conoscenze anzitutto linguistiche, ma anche sul mondo reale della pubblicità, delle vendite e del lavoro. Dobbiamo inoltre fornire al programma procedure per l'applicazione delle suddette conoscenze a questa determinata storia. Tali procedure fanno di solito affidamento su una sorta di processo di ricerca, in cui le risposte desiderate vengono trovate muovendosi in un certo spazio di risposte possibili e di risultati intermedi.

Perchè il nostro programma sia veloce, è importante che tale processo di ricerca sia il più possibile efficiente, intendendo con ciò che dovrebbe esplorare il minor numero possibile di risposte potenziali prima di pervenire a quelle desiderate. Il modo per conseguire questo risultato è generalmente quello di usare come guida le conoscenze insite nel programma.

I programmi di I.A. richiedono diversi tipi di conoscenza, in funzione dei problemi particolari che sono chiamati a risolvere.

Un programma di risposta a domande sulla storia precedente deve disporre dei seguenti tipi di conoscenza:

- Linguistica: lessico, sintassi e semantica della specifica lingua
- Conoscenza della realtà, comprendente la conoscenza di:
  - quali sono gli obiettivi umani
  - come funzionano le aziende

Uno dei problemi più importanti nell'ideazione di un programma di I.A. non consiste solo nel determinare quale conoscenza deve possedere il programma, ma in che modo tale conoscenza può essere rappresentata. L'importante è scegliere una rappresentazione per cui il ragionamento sugli elementi di conoscenza risulti il più possibile semplice e immediato. Pertanto, per poter decidere sulla struttura da dare alla conoscenza, si deve decidere anche sui tipi di ragionamento da farsi.

Chi debba rispondere a quesiti su una storia come quella esposta deve essere in grado di utilizzare le sue conoscenze per fare i seguenti tipi di ragionamento:

- elaborazione linguistica, sia per acquisire la storia che per generare risposte ben organizzate;
- ragionamento deduttivo, per ricavare conclusioni in assenza di certi fatti;
- ragionamento non-monotonico, per ricavare conclusioni in assenza di certi fatti (il significato di non-monotonico verrà spiegato più avanti).

Nelle due sezioni che seguono, si tratterà in modo un po' più dettagliato di queste forme di ragionamento.

4. Elaborazione linguistica

Consiste necessariamente delle seguenti fasi:

- scomporre la storia in unità lessicali e consultare per ciascuna il dizionario;
- determinare la struttura sintattica di ogni frase della storia;
- costruire una rappresentazione "semantica" di ogni frase;
- integrare le frasi in un'unica struttura coerente.

Fasi analoghe, ma essenzialmente in direzione opposta, sono necessarie per fornire, nella lingua utilizzata nella storia, le risposte alle domande. Ci limiteremo in questa sezione, a esaminare la parte di input del pro-

(*) Si può sottilizzare sul problema filosofico di definire cosa significhi capire una storia. Noi qui diciamo che "capire" è un processo con due input, una storia appunto ed un particolare problema. Un computer avrà capito la storia se mette in atto azioni appropriate rispetto al problema dato, il quale è, in questo caso, rispondere a delle domande.

Fig. 1 - Esempio di ambiguità sintattica.
(Legenda: F = Frase;
SN = Predicato Nominale;
SV = Predicato Verbale;
PP = Predicato Preposizionale)

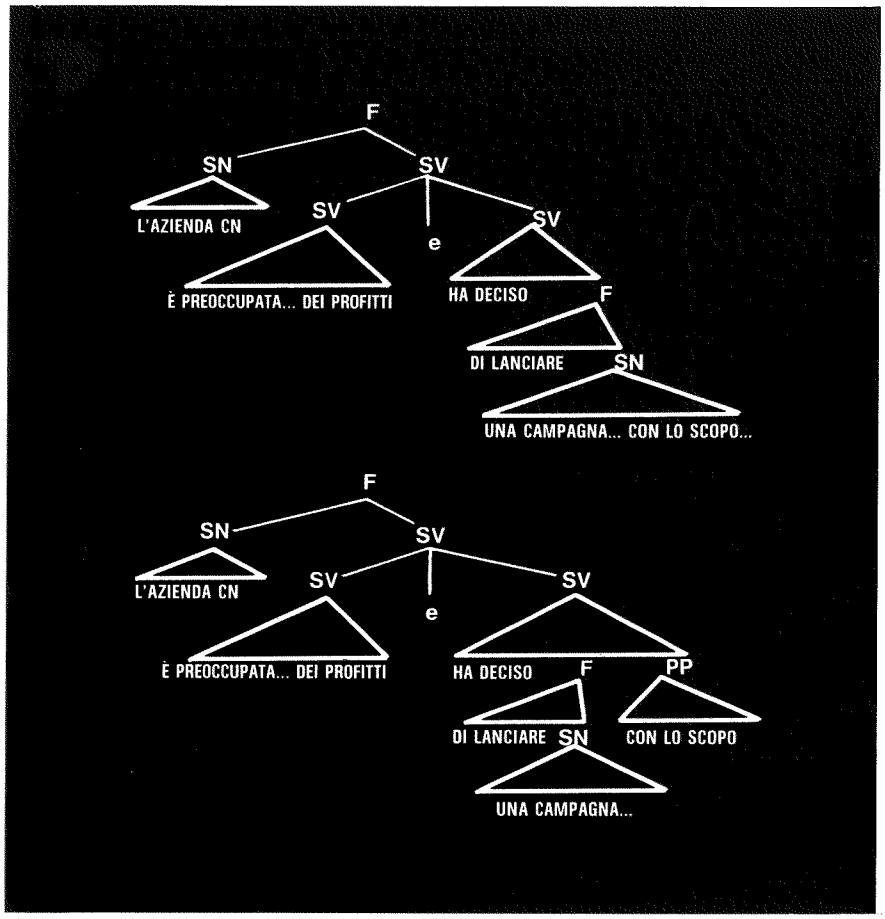
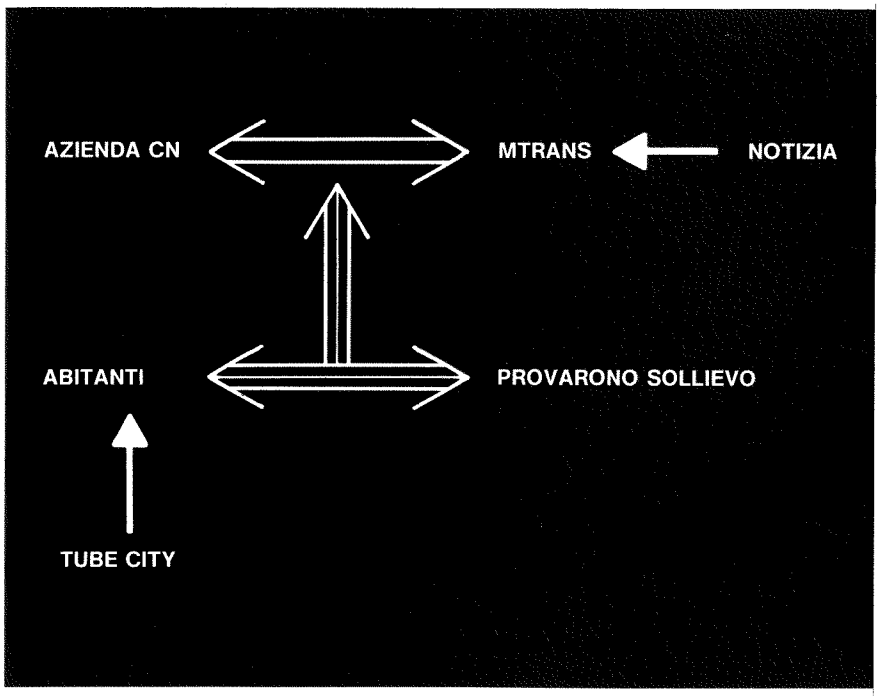


Fig. 2 - Concettualizzazione di una frase.
(MTRANS = Trasferimento Mentale)



cesso. A questo punto è importante notare che, per quanto tutte necessarie, le quattro fasi citate possono anche non essere separate o eseguite nell'ordine indicato. Se si prendono in considerazione i programmi interpretativi del linguaggio naturale che sono attualmente in circolazione [9], si trovano diversi modi per scomporre in parti il processo di interpretazione. L'importante è che tutte queste parti dell'analisi siano fatte, in modo che la stringa iniziale di input venga tradotta in una rappresentazione strutturale delle conoscenze contenute nella storia in questione.

Nessuna parte di questa analisi è irrilevante. Infatti, anche se stiamo discutendo di elaborazione linguistica, non possiamo fare a meno di considerare anche altri tipi di conoscenza. Per capire il perchè, è sufficiente considerare alcuni frammenti della storia. Si consideri ad esempio il problema lessicale presentato dalla proposizione:

«L'annuncio della nuova campagna è giunto come una buona nuova per gli abitanti di Tube City». L'espressione "è giunto" non ha il significato solito di arrivo fisico, ma ha un significato idiomatico. La frase "è giunto" può essere sostituita in questo brano dalla parola "è stato". Anche l'analisi sintattica presenta problemi di ambiguità, sia che venga fatta con una grammatica tradizionale e un sistema parsing, sia con un sistema di accoppiamento basato sul principio delle maschere.

La figura 1 rappresenta due possibili "alberi" di analisi della prima frase della storia. Solo il primo albero ha un senso, poichè un obiettivo non può essere uno strumento, mentre una compagnia può avere un obiettivo.

Diverse tecniche sono disponibili per la rappresentazione del contenuto semantico di una storia come questa. Per la scelta della tecnica adatta, è importante considerare sia la possibilità di rappresentare tutte le informazioni contenute nella storia, che la facilità di scrittura dei programmi che operano su tale rappresentazione e che ricavano nuove informazioni che possono non essere esplicite nella storia.

La figura 2 mostra un modo di rappresentare la frase: "L'annuncio è giunto come una buona nuova per gli abitanti di Tube City".

La rappresentazione si basa su una particolare teoria semantica chiamata "conceptual dependency" (dipendenza concettuale) [6], secondo la quale si parte da un piccolo insieme di "primitive", o azioni elementari, e vengono definiti i modi con cui possono essere combinate tra loro.

Le singole parole della frase devono allora essere rappresentate come strutture contenenti queste "primitive". MTRANS sta per trasferimento mentale (in contrapposizione al trasferimento fisico come quello che sarebbe usato per rappresentare il significato comune di "è giunto"). L'azienda CN ha compiuto il MTRANS della notizia. Questo MTRANS ha dato sollievo agli abitanti di Tube City. Le rappresentazioni della "dipendenza concettuale" di gran parte delle frasi assomigliano alla forma originaria della frase meno di quanto lo siano molti altri tipi di rappresentazione. Esse sono tuttavia ideali per alcuni tipi di ragionamento, in quanto contengono già informazioni che possono occorrere ma che non erano esplicite nella frase originaria (ad esempio che gli abitanti provarono sollievo).

L'integrazione di tutte le frasi di questa storia in una sola rappresentazione coerente richiede la conoscenza delle aspirazioni umane, del mondo degli affari e della pubblicità.

È necessario, ad esempio, trovare un collegamento tra la prima e la seconda frase. Per far ciò, occorre immaginare che L & C debba essere una agenzia di pubblicità e che stia per condurre la nuova campagna.

Una volta letta la storia e costruita la sua rappresentazione, questa può essere passata a chi deve rispondere alle domande.

In funzione della rappresentazione scelta, il compito di quest'ultimo è già stato più o meno assolto. Per completare il lavoro, occorre cambiare la rappresentazione della storia con informazioni supplementari sul mondo reale.

5. Ragionamento

In questo paragrafo si esamineranno una ad una le domande citate nel paragrafo 3 e si vedrà come si possano elaborare le risposte appropriate.

Si consideri la prima domanda "*Perchè l'azienda CN lancia una nuova campagna pubblicitaria?*" La risposta richiede anzitutto la consapevolezza che CN ha *l'obiettivo* di elevare i profitti. Sappiamo che le persone in stato di ansia vogliono spesso fare qualcosa per uscirne.

Per sapere che l'obiettivo è quello di elevare i profitti (in contrapposizione a quello di diminuirli), occorre sapere che la gente vuole di solito fare più soldi, piuttosto che farne meno. Avendo riconosciuto l'obiettivo, proviamo a scoprire il *piano* che ha CN per realizzarlo. Possiamo provare in due modi diversi. Un modo consiste nel ragionare in base a concetti di base. Questo ragionamento potrebbe procedere come segue:

- profitto = [(prezzo - costo della materia prima) x quantità venduta] — spese fisse.
- Si assuma che (prezzo — costo della materia prima) sia maggiore di 0, a meno che l'articolo in questione non sia etichettato come articolo venduto sottocosto a scopo di lancio. Si tratta di un esempio di *ragionamento non-monotonico,* in quanto la sua conclusione è basata sulla mancanza di una parte specifica di informazione (cioè che l'articolo è venduto sottocosto). Se questa parte di informazione diventasse disponibile in seguito, allora la suddetta conclusione dovrebbe essere ritirata. Pertanto la dimensione dell'insieme di cose conosciute (fatti basilari e conclusioni che ne derivano) non aumenta necessariamente nel tempo in modo monotonico. Il *ragionamento "per difetto",* che abbiamo seguito finora, è uno dei tipi più comuni di ragionamento non-monotonico [4].
- Se le spese fisse sono costanti, allora il profitto aumenta con la quantità venduta.
- La pubblicità fa sì che la conoscenza di un prodotto si allarghi ad un numero maggiore di persone.
- Se più persone vengono a conoscenza di un prodotto e ne hanno bisogno, allora lo compreranno. Quindi aumenterà la quantità venduta.
- Una maggiore pubblicità fa salire tuttavia le spese fisse (che comprendono la pubblicità). Perciò il profitto salirà se le maggiori vendite bilanceranno le maggiori spese.

Pertanto l'aumento della pubblicità può essere una tattica ragionevole per elevare i profitti con l'incremento delle vendite. La chiave della ri-

sposta a questa domanda risulta allora la comprensione delle aspirazioni umane e dei piani possibili per realizzarle [7, 10].

La seconda domanda è *"Perchè Smith pensa che la sua iniziativa avrà successo?"* Per rispondere si deve esaminare ancora il ragionamento che sta dietro la risposta al primo quesito. Dobbiamo poi esaminare più in dettaglio una sua fase. Il ragionamento da farsi è il seguente:

- Se altre persone vengono a conoscenza di un prodotto e se esse necessitano di quel prodotto, allora lo acquisteranno. Quindi aumenterà la quantità venduta. Questo è il risultato del ragionamento precedente).
- Se si ha e si usa un oggetto X, si ha bisogno anche di parti di X.
- Gli accumulatori e le candele sono parti dei veicoli.
- La gente ha bisogno di accumulatori e candele.

La pubblicità farà perciò aumentare le vendite di accumulatori e di candele.

Finora abbiamo basato il ragionamento su una conoscenza generale delle vendite. Esistono programmi in grado di fare ragionamenti essenzialmente simili a questo.

In realtà le persone non ragionano così nella maggior parte dei casi. Di solito le persone risolvono i problemi sulla base di quelli che hanno affrontato prima e per i quali hanno già soluzioni memorizzate. Nel caso del problema in questione, si potrebbe ad esempio seguire la regola seguente:

Se: obiettivo = maggiori profitti *e* X è uno dei vostri prodotti *e* il mercato di X è florido

Allora: azione = maggiore pubblicità di X

Insiemi di regole come questa costituiscono la base del comportamento degli esperti in diversi campi. Usando queste regole, è possibile costruire programmi per simulare tali comportamenti [2].

Quando si dispone di regole, i problemi possono essere risolti più rapidamente di quanto sarebbe possibile con un ragionamento derivante dalle conoscenze di base.

Per rispondere al terzo quesito, *"Perchè è insoddisfatto il management di CN?"* sono necessarie due azioni.

Riconosciamo anzitutto che, con la frase "L'azienda CN è preoccupata" intendevamo dire in realtà che il suo management è preoccupato. Le aziende non possono essere preoccupate, solo gli esseri umani possono esserlo. Dobbiamo poi usare la nostra conoscenza del fatto che le persone hanno spesso come obiettivo il far soldi. È inoltre necessario avere una certa conoscenza delle aziende. Possiamo allora ragionare nel modo seguente:

- Obiettivo = guadagnare.
- Negli affari, se non si prendono iniziative, salvo qualche evento fortunato del tutto casuale, si tende a guadagnare di meno.
- Perciò se non siete soddisfatti della situazione attuale, dovete introdurre dei cambiamenti per migliorarla.

Il quarto quesito era *"Perchè gli abitanti di Tube City sono contenti della nuova campagna pubblicitaria?"*

Come per le altre domande, anche la risposta a questa non è immediata. Dopo tutto, gli abitanti di Tube City si occupano di produzione, non di pubblicità. Per rispondere alla domanda, bisogna ragionare così:

- La gente vuole posti di lavoro.
- Se diminuisce la produzione di X, meno persone lavoreranno per la produzione di X.
- Una maggiore pubblicità farà aumentare le vendite.

Possiamo pertanto concludere che una maggiore pubblicità porterà a più posti di lavoro e quindi a buone prospettive per gli abitanti di Tube City.

Quanto detto finora può servire come traccia di ragionamento per chi deve rispondere ai nostri quesiti.

E come si avevano diversi modi di rappresentare la conoscenza della storia, così ci sono rispettivamente diversi modi per realizzare i processi di ragionamento richiesti. Un approccio consiste nel rappresentare la conoscenza sotto forma di formule logiche e di usare poi una procedura formale di prova per portare avanti il ragionamento [1].

Un altro approccio consiste nel rappresentare la conoscenza sotto forma di regole del tipo "se-allora", e di usare poi un metodo di accoppiamento per decidere quando si può applicare una regola. Esiste anche un terzo approccio che consiste nel rappresentare la conoscenza mediante un grafo di oggetti correlati (come il grafo di dipendenza concettuale usato in precedenza), e di usare poi un insieme di procedure che percorrono il grafo e cambiano varie parti tra loro per produrre nuovi risultati. Alcuni programmi usano combinazioni di due o più di queste tecniche [8].

Una considerazione importante, riguardo a queste tecniche, è che conoscenze estranee non devono interferire con la soluzione di un problema specifico.

Le tecniche di ragionamento diretto o immediato considerano spesso diverse modalità per mettere insieme pezzi separati di conoscenze allo scopo di produrre delle conclusioni. Per un determinato problema, queste conclusioni sono tuttavia il più delle volte non significative. Ad esempio, un programma ricco di conoscenze sufficienti per rispondere alle nostre domande sul caso dell'azienda CN, sarebbe anche in grado di concludere che la domanda di pneumatici è in aumento, così come quella di autostrade.

Non occorre tuttavia preoccuparsi di questo per rispondere alle domande rivolte al programma. Pertanto, parallelamente allo sviluppo di tecniche per la rappresentazione della conoscenza e per ragionarci sopra, dobbiamo anche trovare i mezzi per concentrare l'attenzione del programma nei punti giusti. Si stanno facendo progressi in questa direzione. Ad esempio, Eurisko [3] è un programma che scopre idee in vari campi, inclusa la teoria dei numeri interi, in cui ha scoperto, ad esempio, il concetto di numero primo.

Il fattore determinante di questa scoperta è stata la capacità di focalizzare l'attenzione del programma sui percorsi apparentemente più suscettibili di condurre a risultati inte-

ressanti, anzichè su percorsi apparentemente senza sbocco (quelli, ad esempio, che conducevano all'insieme dei numeri con esattamente tredici divisori). In effetti, il modo con cui si compie questa focalizzare è lo stesso con il quale si possono fare molte altre cose con l'I.A., e cioè l'applicazione di più conoscenze. Nel caso portato ad esempio, il programma sapeva che alcuni valori limite (ad es. 0,1,2) rivestono più interesse di gran parte degli altri valori.

6. Conclusione

L'obiettivo dell'intelligenza artificiale consiste nella costruzione di programmi in grado di risolvere problemi difficili.

L'unico modo conosciuto di risolvere tali problemi consiste nel avere la maggior conoscenza possibile nei campi di pertinenza del problema e nella capacità di applicare le conoscenze giuste al momento giusto. Perciò la possibilità di progresso dell'I.A. è legata necessariamente all'acquisizione di ulteriori conoscenze (sul linguaggio, sugli esseri umani e sul mondo reale) ed allo sviluppo di tecniche per sfruttare efficacemente tali conoscenze. Anche gli studiosi di discipline umanistiche cercano di conoscere di più riguardo a molte di queste stesse cose. Da ciò deriva appunto il potenziale di interazione tra le due discipline.

L'intelligenza artificiale ha già percorso una parte del cammino verso il suo obiettivo. Esistono programmi in grado di leggere alcune storie simili alla nostra e di rispondere a quesiti in merito. Non esistono invece programmi in grado di leggere tutte le storie possibili, su ogni argomento, e di rispondere alle relative domande.

Nessun programma ha infatti una base di conoscenze abbastanza vasta. Alcuni programmi sfruttano le tecniche di ragionamento da noi citate e sono in grado di risolvere problemi tecnici in settori specifici come la medicina e l'ingegneria. Ma, di nuovo, non c'è un programma capace di spaziare su tutti questi settori. Molti pezzi esistono già, ma ne mancano molti altri. Dobbiamo dunque trovare il modo per combinare efficacemente i vari pezzi in modo tale che, per ogni specifico problema, si debbano considerare solo le conoscenze che lo riguardano.

7. Bibliografia

[1] Chang, C.L. & R.C. Lee. *Symbolic Logic and Mechanical Theorem Proving.* Academic Press, New York, 1973.

[2] Hayes-Roth, F et al. *Building Expert Systems.* Addison-Wesley, Reading, Mass., 1983.

[3] Lenat, D.B. Eurisko: A Program That Learns New Heuristics and Domain Concepts. The Nature of Heuristic II: Program Design and Results. *Artificial Intelligence,* Mar., 1983.

[4] Reiter, R. A Logic for Default Reasoning. *Artificial Intelligence 13,* Apr., 1980.

[5] Rich, E.A. *Artificial Intelligence.* McGraw-Hill, New York, 1983.

[6] Schank, R.C. Identification of Conceptualizations Underlying Natural Language. In R.C. Schank & K.M. Colby (editors), *Computer Models of Thought and Language,* Freeman, San Francisco, 1973.

[7] Schank, R.C. & R.P. Abelson, *Scripts, Plans, Goals, and Understanding.* Erlbaum, Hillsdale, N.J., 1977.

[8] Simmons, R.F. *Computations from the English,* Prentice-Hall, Englewood Cliffs, N.J., 1984.

[9] Tennant, H. *Natural Language Processing.* Petrocelli, New Xork, 1981.

[10] Wilensky, R. PAM. In R.C. Schank & C.K. Riesbeck (editors), *Inside Computer Understanding.* Erlbaum, Hillsdale, N.J., 1981.

[11] Winston, P.H. & B. Horn. *LISP.* Addison-Wesley, Reading, Mass., 1984.

Tecnologie informatiche per i disabili

TOM HILL, CHARLES HUNT, ROY YEATER

Honeywell
Enabling Technology Task Team
Minneapolis (USA)

1. Introduzione

Questo articolo costituisce una documentazione parziale delle attività svolte da tecnici della Honeywell per trovare soluzioni tecnologiche a problemi di portatori di handicap.

Non occorre sottolineare la necessità di questo tipo di progetti. Solo negli U.S.A., 21 milioni di cittadini soffrono di minorazioni fisiche di vario genere, che ne fanno una delle minoranze più consistenti degli Stati Uniti (e si tratta dell'unica minoranza cui può appartenere chiunque). Dodici milioni di loro sono in età lavorativa. Il tasso di disoccupazione delle persone affette da minorazioni fisiche supera il 60%, è cioè quasi nove volte superiore al tasso nazionale.

Abbiamo citato questi numeri per porre nella giusta prospettiva questo problema, benchè progetti come quelli di cui parliamo non trovino davvero una giustificazione soltanto nei numeri.

Se infatti ci fossero anche solo poche persone in stato di necessità, si dovrebbe comunque aiutarle.

Due dei tre progetti qui descritti sono rivolti alle necessità di una piccola parte della popolazione affetta da handicap, cioè alle persone con minorazioni gravi. Esse possono beneficiare più di ogni altri della tecnologia, benchè si debba osservare che anche handicap apparentemente minori esigono attenzione. Ad esempio, la quadriplegia impedisce l'uso di una tastiera standard, ma la stessa cosa vale per una artrite leggera.

Dobbiamo tenere presente inoltre che ogni progetto rappresenta solo un passo verso l'obiettivo di consentire ai disabili di trovare un impiego. Ad esempio, due progetti, EYECOM e COMMIS, si occupano del problema di accesso alla tastiera. Un minorato fisico, che utilizzi uno di tali progetti, si troverebbe tuttavia ancora di fronte ad altri problemi per poter lavorare come programmatore. Possiamo citare ad esempio uno degli ostacoli più semplici: l'impossibilità di sollevare e maneggiare i manuali di software.

Ciò non significa che i due progetti siano inutili.

Ognuno di essi è potenzialmente importante, ma ciascuno offre una soluzione solo ad una parte di un problema più vasto.

Molti problemi impegnativi attendono una soluzione, e lo "Honeywell Enabling Technology Task Team" (Gruppo di lavoro su tecnologie informatiche per i disabili) ne ha stilato un non breve elenco.

Nel seguito vengono descritti tre dei progetti realizzati, ai quali hanno partecipato gli estensori dell'articolo. Ma molte altre persone hanno dedicato volontariamente il loro tempo e le loro capacità tecniche per contribuire alla realizzazione di queste apparecchiature.

2. Il Progetto EYECOM: comunicare con gli occhi

Il Progetto EYECOM (EYE COMmunication) ha come obiettivo lo sviluppo di un sistema di comunicazione destinato a chi è in grado di comunicare ("parlare") solo con gli occhi.

La nostra aspirazione è di fornire, con questo sistema, a persone affette da mutismo o con un controllo motorio estremamente limitato, la capacità di comunicare con relativa facilità e, come obiettivo finale, la possibilità si svolgere un lavoro normale.

Il fulcro del sistema è un dispositivo sviluppato al Denver Research Institute della Università di Denver da George Rinard, direttore del progetto, Owen Barnes e Paul Manning. Il dispositivo consiste essenzialmente in un microprocessore pilotato da un sensore oculare, o trasduttore. Il trasduttore individua la direzione dello sguardo dalla posizione degli occhi riferita alla testa. In combinazione col microprocessore, esso consente all'utilizzatore di indicare un simbolo presente su uno schermo video, fissando semplicemente il quadrante appropriato di tale schermo. Il seguito dipende dal software che interpreta i simboli e dalle unità periferiche con cui comunica il microprocessore; l'uso del trasduttore permette di comandare la sedia a rotelle, stampare messaggi o programmare un computer.

Nel dicembre 1984 abbiamo acquistato un apparecchio a controllo oculare dal Denver Research Institute. Da allora abbiamo sviluppato software per l'adattamento dell'apparecchio alle necessità di invalidi del "Courage Center" a Golden Valley nel Minnesota.

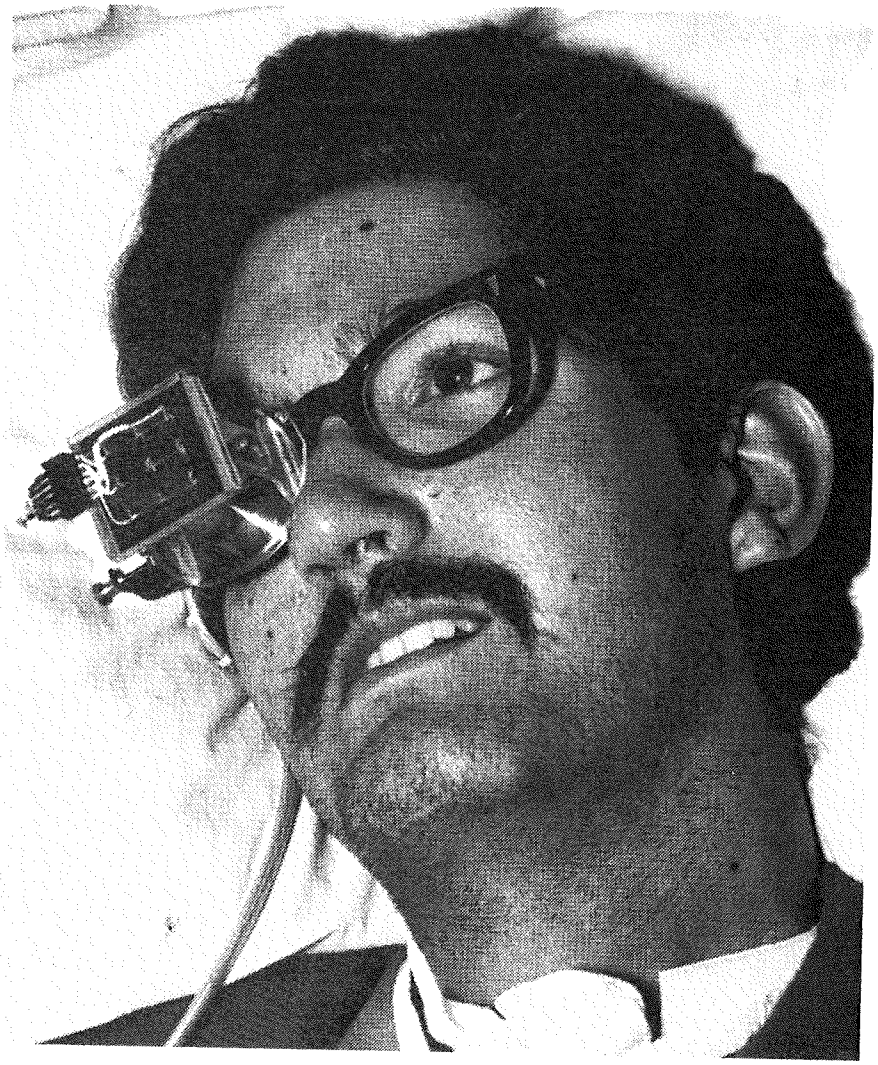


Fig. 1 - Il sistema EYECOM consente a persone aventi soltanto il controllo dei movimenti oculari, di comunicare e anche di svolgere un lavoro, come programmare. Il sistema si basa su un particolare dispositivo che rileva la direzione dello sguardo.
In figura è mostrato lo speciale paio di occhiali usato. Al posto della lente di destra è montato il dispositivo attraverso cui l'utente vede un display con 64 caselle, contenente altrettanti simboli. Egli può scegliere uno di tali simboli semplicemente fissandovi lo sguardo. Può così comporre parole e frasi che, attraverso l'altro occhio, vede man mano visualizzate sullo schermo di un monitor o stampate sulla carta.

Il trasduttore consiste in un diodo che emette raggi infrarossi (LED), montato sul ponte di un paio di occhiali, uno specchietto per raggi infrarossi, o "specchietto caldo", montato al posto della lente destra, una lente ed un chip di memoria dinamica ad accesso casuale (RAM), montati sulla stanghetta degli occhiali. I raggi provenienti dal LED sono riflessi dallo specchietto verso la cornea dell'occhio e dalla cornea verso la memoria RAM.

La macchia di luce si sposta con la cornea poichè il raggio di curvatura della stessa è inferiore a quello dell'occhio. La cornea si comporta come la sezione di un piccolissimo specchio sferico inglobato in una sfera più grande.

Quando l'occhio si muove, la cornea ruota attorno al centro dell'occhio, non attorno al proprio centro. A causa del piccolo movimento della cornea, si spostano anche i raggi riflessi. Se la cornea avesse la stessa curvatura dell'occhio, i raggi riflessi rimarrebbero invariati, esattamente come i raggi riflessi da uno specchio sferico rotante.

Il sensore è una RAM dinamica, cioè una matrice di celle di memoria che devono essere rigenerate a intervalli di pochi millisecondi. Le celle con carica elettrica superiore ad una soglia predeterminata sono lette come "1" logico, mentre quelle con tensione inferiore sono lette come "0" logico. Se le celle non vengono rigenerate periodicamente, portatori di carica eccitati termicamente provocano la dispersione graduale della carica elettrica immagazzinata nelle celle, modificando l'"1" logico in "0" logico. È appunto questa caratteristica che permette l'uso del chip come sensore; la carica delle celle esposte alla luce si attenua molto più rapidamente della carica di quelle non esposte.

La posizione del raggio riflesso sulla matrice di celle viene determinata in un ciclo a due fasi. La RAM viene inizializzata, registrando "1" logici in tutte le celle. Quindi il LED viene attivato per 45 millisecondi. In questa fase la matrice non viene rigenerata e la carica elettrica delle celle colpite dalla luce si disperde. Nella seconda fase del ciclo, il LED si spegne e il microprocessore comincia a leggere le celle di memoria. Per assicurare il mantenimento della configurazione di cariche registrate dalla luce riflessa, la matrice viene rigenerata durante questa fase ogni due millisecondi. Ogni volta che il microprocessore trova uno "0" logico, registra la relativa posizione di memoria e ricarica la cella. Pertanto alla fine della scansione della matrice, tutte le celle conterranno "1" logici e la prima fase del ciclo può essere ripetuta.

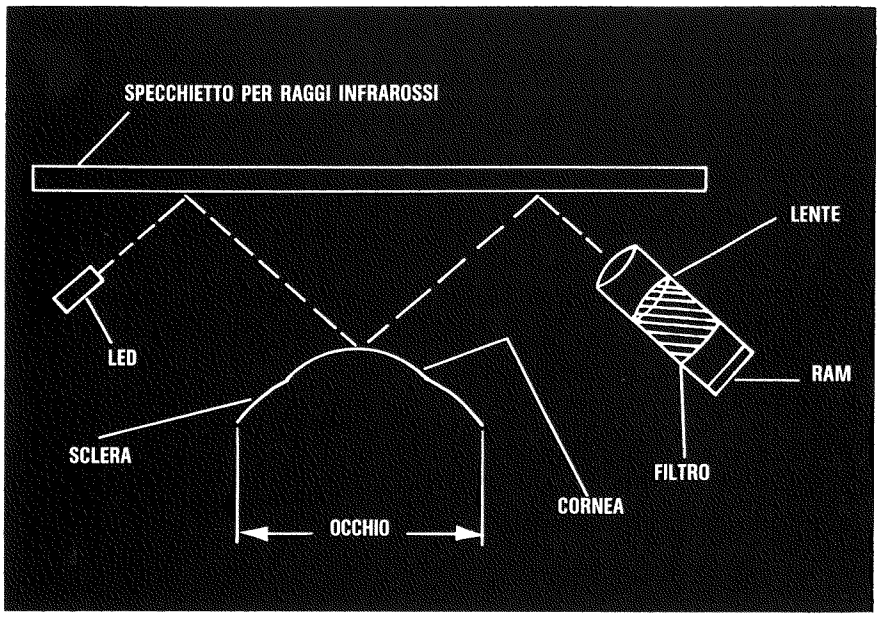


Fig. 2 - Il sistema di rilevamento dei movimenti oculari consiste di un diodo ad emissione di raggi infrarossi (LED), uno specchietto per raggi infrarossi, una lente con filtro ed un chip di memoria random-access (RAM). La posizione della cornea determina l'ampiezza dell'angolo di riflessione della luce e quindi la posizione sul chip di memoria della macchia luminosa riflessa. Prima di ogni ciclo di operazione, ogni cella di memoria riceve una carica elettrica. Questa carica si disperde più rapidamente nelle celle su cui incide la luce, che non nelle celle vicine. Pertanto la lettura della memoria permette di rilevare la posizione della macchia luminosa riflessa. Un filtro applicato davanti al chip assorbe la luce visibile e, in combinazione con lo specchietto per raggi infrarossi, blocca quasi completamente l'arrivo sul chip di luce emessa da sorgenti diverse dal LED.

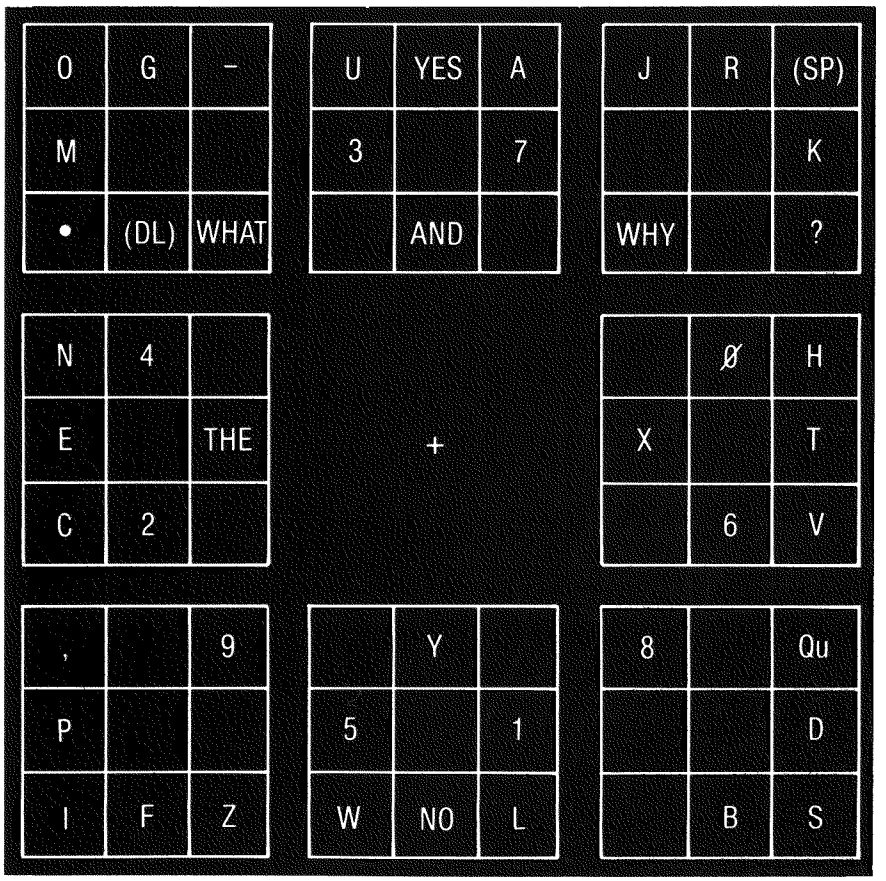


Fig. 3 - Il display qui mostrato consente di scegliere uno tra un rilevante numero di caratteri, per mezzo di pochi movimenti dell'occhio. Il centro del display è la posizione di riferimento. Per scegliere, ad esempio, la lettera H, l'utilizzatore fissa lo sguardo al centro, poi sulla casella centrale a destra, e quindi sul quadrante superiore a destra, tornando infine al centro. I movimenti più semplici corrispondono ai caratteri ed alle parole usate più frequentemente.
Si tratta dei caratteri E e O, la parola YES, lo spazio, i caratteri T e S, la parola NO ed il carattere I. Questi simboli possono essere scelti semplicemente guardando l'apposito quadrante del display e quindi al centro. La casella centrale vuota di ogni quadrante non può essere usata, ma le altre caselle vuote potrebbero contenere parole suggerite dall'utilizzatore. Uno spazio potrebbe anche essere usato per inserire un simbolo di controllo che permetta all'utente di passare ad un altro insieme di 64 caratteri.

Durante la prima fase del ciclo il microprocessore è disponibile per trattare l'informazione ricevuta nel ciclo precedente.

Mentre la cella di memoria registra la riflessione successiva, il microprocessore determina il simbolo cui corrispondono le coordinate x, y dell'ultima configurazione di "0" logici.

L'intero ciclo impiega 80 millisecondi; ci potrebbero essere teoricamente fino a 12 selezioni al secondo.

I caratteri vengono scelti con un metodo chiamato ETRAN, sviluppato dal Jack Eichler. Il display mostrato nella figura 3 è montato quattro centimetri circa davanti allo specchietto, che è trasparente ai raggi infrarossi. Una lente di Fresnel posta tra il display e lo specchietto consente all'utilizzatore la messa a fuoco ad una distanza così piccola.

Il centro del display è la posizione di riferimento da cui si misurano i movimenti. Per scegliere un carattere, l'utente fissa lo sguardo al centro, poi sulla casella che contiene il simbolo desiderato. Quindi guarda l'area del display che corrisponde al-

la posizione del carattere nella casella già prescelta. Per scegliere ad esempio una N, egli fissa lo sguardo al centro del display, poi sulla casella al centro della riga a sinistra, quindi sul quadrante superiore sinistro, tornando infine al centro del display. La posizione centrale è l'ultima della sequenza di movimenti dell'occhio ed indica che si è conclusa la scelta del carattere.

Il sistema con cui stiamo lavorando offre diversi tipi di retro-azione. Un LED visibile è posto al centro di ciascuna casella del display. Quando l'utente sceglie un carattere, si accende il LED posto nella corrispondente casella. Il microprocessore è collegato al sintetizzatore di voce Speak-N-Spell prodotto dalla Texas Instruments, che pronuncia il carattere scelto. Può anche essere collegato ad una stampante, ad un computer o ad altre unità periferiche.

Poichè la RAM viene gestita dal software, risulta relativamente facile ovviare a problemi connessi con la selezione dei caratteri. Ad esempio, l'algoritmo che determina la posizione dell'occhio dall'output della RAM, può essere adattato per compensare un limitato movimento dell'occhio.

Tra i nostri obiettivi c'è anche il miglioramento dell'algoritmo per l'auto-taratura del sistema, in modo che il sistema possa adattarsi automaticamente quando gli occhi dell'utente lacrimano o gli occhiali scivolano.

Lo scopo principale del nostro progetto è tuttavia quello di adattare il sistema per il collegamento ad un personal computer in modo da mettere a disposizione la sua potenza a coloro che non sono in grado di usare la tastiera. Stiamo attualmente revisionando il display ed il software in modo da consentire la selezione di caratteri speciali, e espandendo l'interfaccia RS-232C in modo che possa consentire il trasferimento al computer di tutti i comandi standard della tastiera. Entro poco tempo dovrebbe essere possibile agli utenti del sistema a controllo oculare di eseguire programmi standard su un computer Apple.

3. Un centralino telefonico a comando vocale

Il riconoscimento della voce si applica comunemente ad attività che sembrano richiedere più di due mani. Per fare un esempio, un sistema con ingresso dati vocale permette ad un addetto ai controlli di qualità di applicare dei dati a oggetti e contemporaneamente di manipolarli. Il Technology Strategy Center della Honeywell (TSC) ha utilizzato il riconoscimento della voce per uno scopo diverso: per consentire a persone con uso limitato o nullo delle mani, di eseguire un compito che è di norma manuale. Il compito è quello di usare il telefono ed il sistema consiste di un computer controllato da comandi vocali, che esegue azioni di tipo manuale, come sollevare un ricevitore o comporre un numero. Noi contiamo che questo sistema possa permettere ai disabili di essere impiegati come operatori di centralini o in altre attività nell'ambito telefonico.

Il progetto ebbe origine in occasione della visita al "Courage Center" di Minneapolis di un gruppo di ingegneri del TSC della Honeywell. Dopo questa visita, discutemmo le tecnologie di cui esiste competenza al TSC e che potrebbero essere applicate per aiutare i disabili. Fra i più entusiasti a lavorare con noi fu Ray Fulford del Courage Center, il cui problema principale era quello di trovare un impiego per i quadriplegici. Fummo così indotti a scegliere il riconoscimento della voce come la tecnologia su cui puntare in quanto i quadriplegici non sono normalmente affetti da deficit della parola. Il progetto fu finalizzato verso l'uso di un PBX (centralino telefonico privato), poichè ritenemmo che le apparecchiature necessarie per questo lavoro potessero aprire la strada verso altre applicazioni, ad esempio al lavoro di addetto alla reception o simili. In una fase successiva del progetto siamo riusciti ad individuare un lavoro specifico: quello di operatore presso la "Honeywell Voice Network" (Rete di Trasmissione Vocale, che collega tutte le sedi Honeywell in USA).

Il sistema, costituito con hardware di facile reperibilità commerciale è costituito essenzialmente da un personal computer, programmato per l'interpretazione di input vocale ed il controllo di un telefono tramite una interfaccia programmabile.

L'interfaccia dell'utente verso il programma del computer è un riconoscitore di voce tarato sull'operatore, che dipende cioè dall'addestramento da parte dell'operatore stesso. Per istruire il sistema, l'utente pronuncia le parole o frasi che sono nel vocabolario del sistema. Questo registra e memorizza la configurazione, o maschera, di ciascuna frase. Nella fase operativa, il sistema confronta le frasi pronunciate dall'utente con le maschere. Se c'è corrispondenza, il codice della frase in questione viene trasferito nel programma del computer. Il vocabolario del sistema consiste di 15 comandi di sistema e dei numeri da zero a nove.

L'utente parla attraverso un microfono a cuffia sia al riconoscitore di voce che con le persone che chiamano. Questo congegno gli consente di conversare con chi chiama e nel medesimo tempo di controllare il telefono. Il computer visualizza le informazioni sullo stato del sistema, come "pronto a rispondere alla chiamata" oppure "richiedente in linea".

Visualizza inoltre il vocabolario valido nello stato attuale del sistema e, su comando, informazioni sulla rubrica telefonica.

Come funziona il sistema? Supponiamo che il telefono suoni. L'operatore dice "Sgancia" e il computer sgancia davvero il ricevitore, connettendo la chiamata. Se chi chiama chiede di trasferire il collegamento su un altro numero, l'operatore risponde "Sto trasferendo, resti in linea per favore". La frase "resti in linea" è un comando che dice al computer di porre il richiedente in attesa. Mentre questi attende, l'operatore emette altri comandi vocali che fanno comporre il numero richiesto e completare il trasferimento da parte del computer.

Il vocabolario del sistema è stato scelto con cura, onde consentire simultaneamente la conversazione ed il controllo del telefono. Le parole e frasi di comando sono quelle nor-

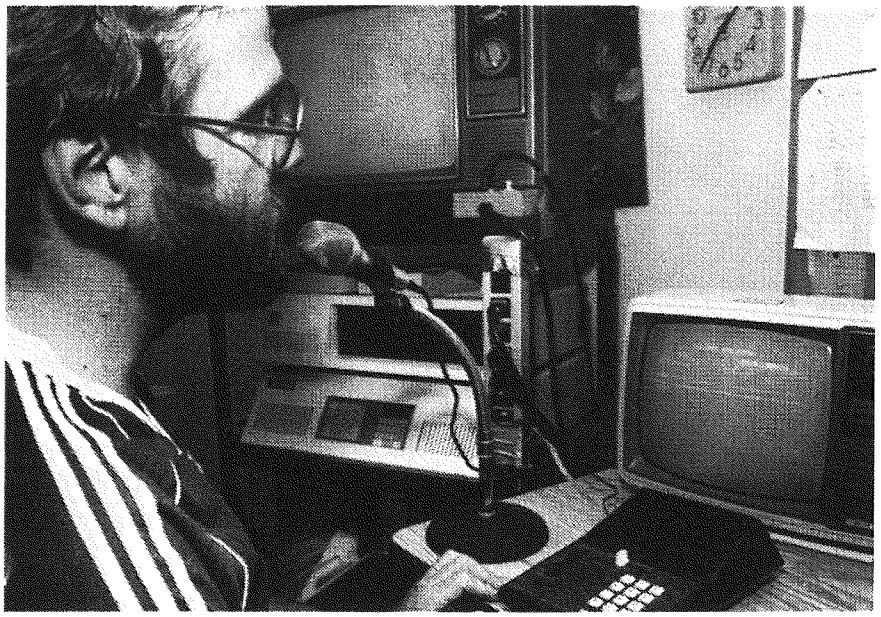


Fig. 4 - L'interfaccia a comando vocale applicata al sistema ROADNET consente all'operatore di rispondere a chiamate, porre l'utente in attesa, comporre numeri telefonici, il tutto con la sola voce. L'interfaccia traduce infatti in azioni manuali i comandi verbali dell'operatore.

malmente usate nella conversazione, così che chi chiama non dovrebbe rendersi conto di alcunchè di diverso nel corso della transazione. Benchè l'alternarsi di conversazione e di comandi non abbia creato problemi durante le prove del sistema, qualsiasi dispositivo di riconoscimento della voce rimane tuttavia soggetto ad errori. Per questo motivo abbiamo previsto dei comandi che consentono all'utente di rimediare a manovre sbagliate o ad errori di riconoscimento, salvo che a sconnessioni accidentali.

Per garantirsi contro sconnessioni accidentali, il comando di riaggancio del ricevitore è stato modificato in modo da accettare una frase multisillabica complessa, come "good bye now", anzichè una sola parola, come "bye". Infatti il sistema è meno suscettibile a confondere "good bye now" con altre frasi. Siamo stati infine in grado di sfruttare una caratteristica del riconoscitore per ridurre ulteriormente il numero di sconnessioni accidentali. Il sistema è adatto a riconoscere il parlato come costituito da parti separate anzichè come un continuo; in altri termini, chi parla deve effettuare una pausa prima e dopo ciascuna frase. Se una delle due pause manca, il sistema non tenta l'accoppiamento della frase con le maschere, evitando così la possibilità di errori. Con un po' di pratica, l'utente può imparare a conversare con pause sufficientemente lunghe e può così ridurre il numero di tentativi di accoppiamento inopportuni.

Una dimostrazione del Centralino telefonico a Controllo Vocale è stata eseguita sul PBX GTE OMNI-S-II. Il successo iniziale è stato incoraggiante, ma ci siamo resi conto di avere in questo modo provato soltanto la validità del concetto, e che dovevano ancora verificare il sistema in un effettivo ambiente di lavoro. Prima di arrivare a ciò, si doveva ancora risolvere una serie di problemi.

Ad esempio, proprio per la nostra ipotesi di partenza che l'operatore avesse un uso limitato delle mani, dovevamo fornigli il mezzo per accendere il sistema e istruirlo a riconoscere la sua voce. Gli interruttori di potenza del computer si trovano spesso in posizioni scomode anche per persone perfettamente normali, ed inoltre per istruire il sistema al riconoscimento della voce è necessario usare la tastiera.

Abbiamo così adattato il sistema sperimentale ad operare col sistema ROADNET della Honeywell.

La Honeywell impiega già parecchi disabili come operatori fuori-orario del ROADNET. Essi, lavorando in casa propria, assistono gli utenti nell'istradamento delle chiamate.

Per adattare il sistema a questo lavoro, abbiamo aggiunto un programma che guida l'operatore nel compito di istruire il sistema. Il computer visualizza la parola del vocabolario da pronunciare e chiede all'operatore di pronunciarla al momento giusto. Il programma consente all'operatore di istruire il sistema senza dover usare la tastiera e di memorizzare su un floppy disk le maschere della sua voce, per cui egli non deve istruire nuovamente il sistema ogni volta che lo usa.

Anche l'hardware del sistema ha dovuto subire lievi modifiche per adattarsi all'applicazione. L'operatore di ROADNET usa spesso l'equivalente funzione vocale del disco combinatore, e molto meno le funzioni di trasferimento chiamata o di messa in attesa. In queste circostanze, la cuffia non è l'ideale, sia perchè c'è solo una limitata sovrapposizione tra ciò che l'operatore deve dire a chi chiama ed i comandi che egli deve dare al sistema, sia perchè chi chiama è in condizione di udirlo mentre dà le cifre per comporre un numero. Abbiamo allora sostituito la cuffia con uno speakerphone ed un microfono; l'operatore comunicherà con l'utente tramite speakerphone e con il sistema di riconoscimento della voce tramite microfono. Quando desidera inviare comandi al sistema, preme un pulsante che esclude lo speakerphone.

Il sistema così modificato sta ora per diventare operativo, grazie anche all'adesione ed al supporto da parte dei responsabili tecnici ed amministrativi del sistema ROADNET del

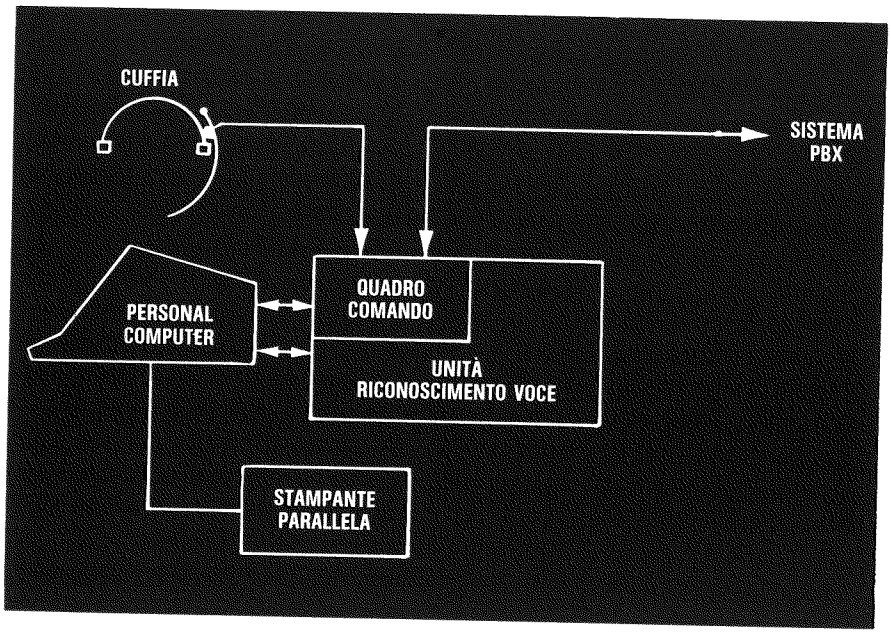


Fig. 5 - La figura mostra la configurazione hardware della interfaccia. Il microfono della cuffia, con l'ausilio di uno speciale vocabolario di comandi, consente all'operatore di controllare il sistema e contemporaneamente di conversare con chi chiama. I comandi vocali sono ricevuti, interpretati e trascodificati dall'unità di riconoscimento della voce. L'uscita di questa unità è inviata al computer, che gestisce il quadro di comando e traduce in azioni i comandi ricevuti.

Corporate Information Management della Honeywell. Siamo convinti che il sistema sarà di ausilio per l'operatore ROADNET, e si dimostrerà un esecutore affidabile e soprattutto un mezzo efficace per aprire nuovi posti di lavoro per i disabili.

4. COMMIS: Interazione senza tastiera con un computer

Il sistema Commis è una interfaccia di computer che, a sviluppo ultimato, consentirà lo svolgimento di programmi applicativi su computer Apple o di altra marca, da parte di persone che non possono usare una tastiera.
Il sistema consiste di un interfaccia software d'utente, di un computer d'interfaccia e di un computer applicativo. Il computer d'interfaccia può anche essere utilizzato in modo autonomo, consentendo all'utente di costruire e stampare un testo.

Commis era all'origine un progetto limitato, finalizzato alla scrittura di un programma che consentisse, usando un manipolatore tipo "joystick", la scansione dell'alfabeto e la scelta di caratteri, costruendo così un testo.

Il progetto attirò invece un interesse maggiore del previsto. Fu applicato alla Como Elementary School di St. Paul nel Minnesota, che gestisce i progetti educativi speciali per la città, e al Gillette Hospital center di St. Paul, l'ospedale regionale per bambini e giovani affetti da minorazioni fisiche. Gli insegnanti della Como Elementary School chiesero poi un programma di tipo più generalizzato, che mettesse in grado i bambini di risolvere ad esempio problemi matematici.

Ci si rese conto a questo punto che era necessario fare un passo indietro e rivedere il problema. La prima possibilità cui si pensò, era quella di scrivere speciali programmi applicativi, in cui fosse incorporata la capacità di gestire display speciali.

Un programma di calcolo potrebbe, ad esempio, visualizzare le cifre da 0 a 9, i comuni operatori aritmetici, +, − e = ed alcuni simboli speciali, come C per "Cancella".
Tutto ciò comporterebbe, però, la stesura di nuovi programmi per ogni applicazione e per ogni tipo di computer.

Sarebbe preferibile trovare il modo di sfruttare programmi applicativi esistenti, perchè in tal caso dovremmo scrivere solo programmi di interfaccia per visualizzare i comandi appropriati ai vari programmi applicativi.

Sorge a questo punto la domanda: "È possibile eseguire i programmi di interfaccia ed i programmi applicativi sullo stesso personal computer?" Per eseguire i due programmi, il computer deve disporre anzitutto di un sistema operativo multitasking od equivalente.

Ci sarebbero tuttavia altre due difficoltà. La prima è che molti programmi applicativi richiedono l'intero schermo video. La seconda difficoltà deriva dall'esigenza di trasferire informazioni dal programma di interfaccia al programma applicativo, mentre tali programmi sono predisposti per input da tastiera, non da un altro programma.

La soluzione adottata consiste nell'impiego di due computer: un computer d'interfaccia per elaborare e visualizzare i "menù" dei comandi ed un secondo computer per eseguire e visualizzare i programmi applicativi.

Questa architettura, unita a un dispositivo hardware specializzato, l'emulatore di tastiera, facilita il nostro compito. L'emulatore di tastiera, che collega tra loro i due computer, risolve il problema delle comunicazioni tra i programmi. Riceve l'output seriale ASCII dal computer di interfaccia e lo trascodifica nei codici di tastiera riconosciuti dal computer applicativo. Nell'attesa di poter disporre dell'emulatore di tastiera, usiamo attualmente il sistema Commis in modo "stand-alone".

Commis ha fra l'altro il vantaggio, rispetto a gran parte degli apparecchi per l'accesso ai computer da parte di disabili, di basarsi essenzialmente su software specializzato anzichè su hardware specializzato.

Dei dispositivi hardware, chiamati

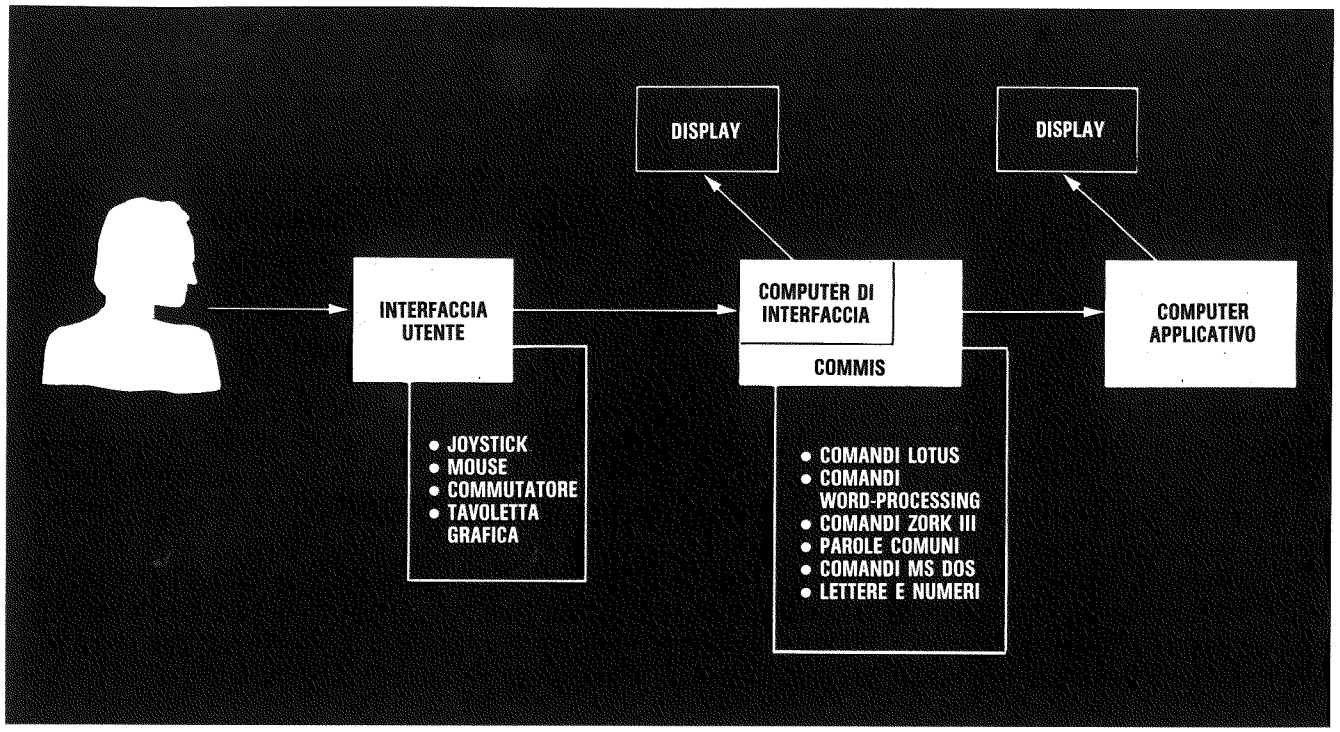


Fig. 6 - COMMIS è una interfaccia che consente l'esecuzione di programmi applicativi standard MS/DOS da parte di persone che non possono usare una tastiera. Il sistema consiste di due computer: un microcomputer che visualizza i menù di comandi appropriati ai vari programmi applicativi, ed un secondo che elabora i programmi applicativi stessi. Il collegamento tra i due computer è realizzato da un emulatore di tastiera, che trascodifica l'output seriale ASCII del computer di interfaccia nel codice della tastiera che è riconosciuto dal computer applicativo. Poichè Commis è un sistema software, può essere facilmente personalizzato secondo le esigenze dell'utente. In figura, sono rappresentati su fondo grigio alcune unità alternative di input e flussi di comandi che possono essere visualizzati del computer di interfaccia.

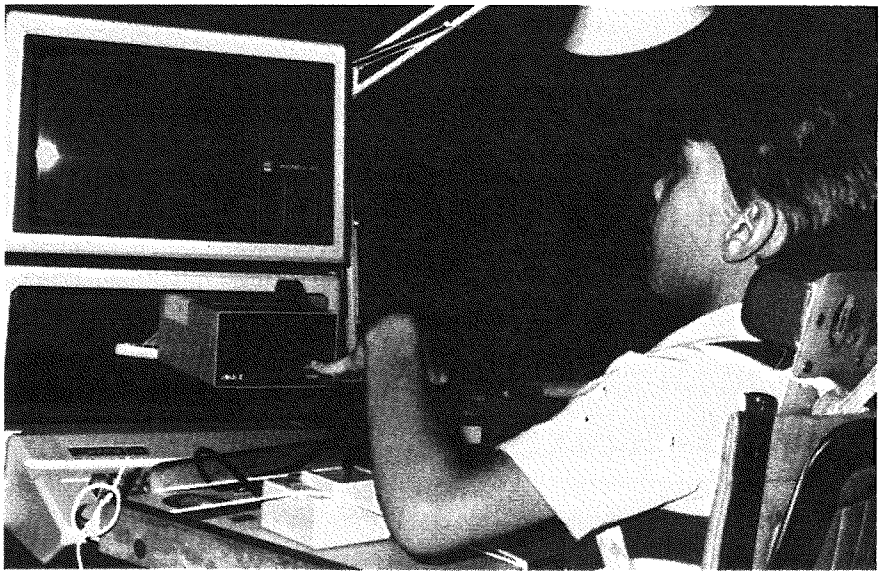


Fig. 7 - Il sistema COMMIS supporta due modi operativi. Il microcomputer che esegue il software Commis può essere usato o come interfaccia verso un altro computer o come unità autonoma. In questo secondo modo di impiego, il sistema permette all'utente di costruire un testo. In questa foto, un piccolo utente compie una selezione di caratteri usando un manipolatore tipo "joystick".

"scanditori", possono essere collegati ad un computer tramite l'emulatore di tastiera, nonchè schede firmware, praticamente con la stessa funzione degli scanditori, per alcuni tipi di personal computer. Gli scanditori esplorano automaticamente l'alfabeto. Caratteri o gruppi di caratteri vengono selezionati in un primo momento (essendo la selezione generalmente un processo in due fasi) da un indicatore LED. Quando si sceglie il carattere desiderato, l'utente lo indica attivando un interruttore o altro dispositivo di input. È superfluo dire che tutti i suddetti dispositivi hardware non sono affatto flessibili. Non consentono ad esempio di modificare i caratteri visualizzati, la velocità di scansione, nè l'unità di input.

È facile, per contro, modificare il sistema Commis.

Esso è predisposto per accettare input da un joystick, ma può essere facilmente adattato per accettare l'in-

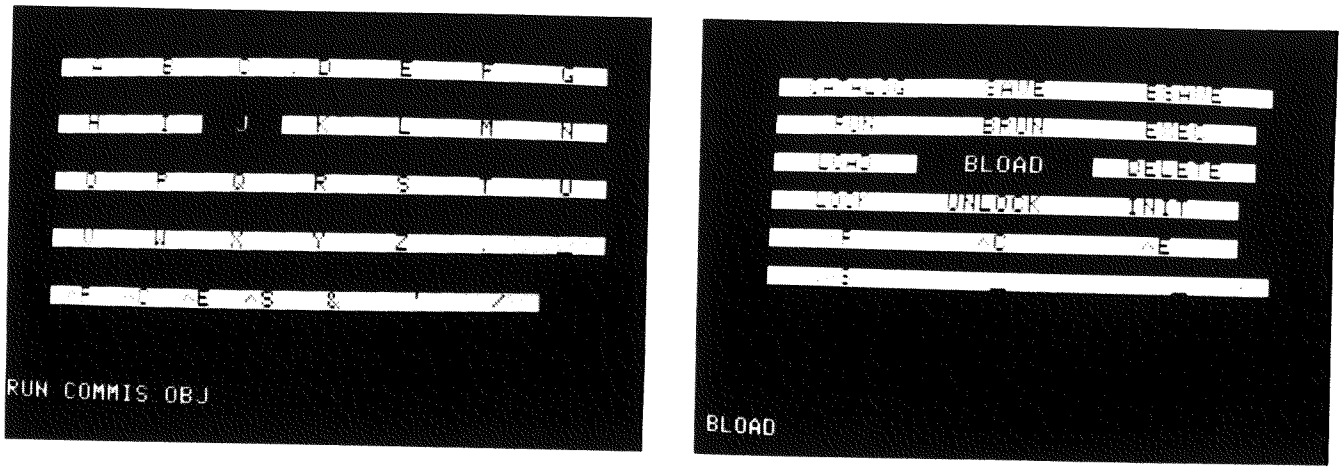


Fig. 8 - I menù bi-dimensionali aumentano il numero di elementi visualizzabili contemporaneamente. Sono qui rappresentati un menù che permette all'utente di costruire un testo mediante selezione di caratteri ed un menù di comandi per Apple.

put da altri dispositivi usati da alcuni disabili per comandare le sedie a rotelle con motore elettrico. Risulta anche relativamente facile modificare alcune caratteristiche del display, quali la velocità di scansione o la durata di attivazione del dispositivo di input necessaria prima di una selezione.

Il costo del sistema Commis risulta inferiore rispetto alle soluzioni hardware. Gli scanditori, ad esempio, sono relativamente costosi, partendo da un prezzo di circa $ 1.500. Possiamo stimare un rapporto di costi di uno a due rispetto ai dispositivi hardware meno costosi e di uno a sei rispetto a quelli più cari.

Una caratteristica di Commis merita un esame più attento: il display flessibile. Commis consente di attrezzare diversi flussi, che visualizzano i comandi appropriati a programmi applicativi diversi.

Un flusso potrebbe, ad esempio, visualizzare i comandi per un programma di "word-processing", un secondo i comandi per un programma di "foglio elettronico" (per la raccolta in tabelle di dati provenienti da archivi diversi) ed un terzo programma i comandi appropriati quando il computer lavora in emulazione di un terminale.

Questi flussi potrebbero essere opportunamente definiti allo scopo di ridurre al minimo il numero di selezioni che un utente disabile è chiamato a compiere.

Un flusso potrebbe, ad esempio, visualizzare il menù delle clausole di indennizzo assicurativo. L'utente può allora, in modo molto semplice, scegliere tra di esse, senza la necessità di inserire la clausola scelta un carattere per volta. Se la clausola desiderata non si trova nel menù, egli potrebbe passare ad un menù di caratteri ed inserire la clausola scegliendoli opportunamente.

Una difficoltà d'impiego dei displays sta nel numero limitato di caratteri che possono contenere, da cui derivano limitazioni nel numero di selezioni possibili e nella lunghezza di ciascuna selezione.

Allo scopo di aumentare il numero di selezioni possibili, abbiamo strutturato alcuni tipi di menù sotto forma di matrici con una entrata in ogni casella, anzichè sotto forma di elenchi con una entrata su ogni riga.

Il sistema si presenta all'utente come un insieme di menù bi-dimensionali.

Permane tuttavia l'impossibilità di mettere sullo schermo più di un certo numero di caratteri. I due formati di flusso che il sistema Commis supporta consentono di ovviare a questa limitazione. Se l'utente lavora con il primo formato, quello che vede ottiene. Se lavora con il secondo formato, il simbolo visualizzato è l'identificatore numerico di una stringa più lunga (fino a 255 caratteri) in cui è trascodificato il simbolo stesso, una volta selezionato. Occorre aggiungere che i due formati possono anche essere frammisti in un unico flusso di visualizzazione.

Il sistema Commis supporta anche un flusso di configurazione utente che consente la personalizzazione del sistema secondo le minorazioni fisiche dell'utente.

Usando questo flusso è possibile, ad esempio, regolare la durata dell'intervallo di tempo tra l'attivazione del dispositivo di input e la registrazione della selezione effettuata.

Una versione preliminare del sistema Commis è stata rivista e verificata dal personale della Como Elementary School e del Gillette Children Hospital e stiamo ora procedendo alla realizzazione della versione definitiva.

Dati sperimentali sullo sviluppo di grandi sistemi di software

ANTONIO PIZZARELLO

Honeywell
Large Computer Products Division
Phoenix, Arizona (USA)

1. Introduzione

Lo sviluppo di grandi sistemi di software rappresenta un grosso sforzo scientifico, tecnico ed organizzativo. Questi sistemi implicano infatti alti livelli di complessità e richiedono il lavoro coordinato di notevoli quantità di specialisti per periodi di parecchi anni.

Esiste quindi un forte interesse a capire quali sono i fenomeni collegati a questi sviluppi, per poterli controllare e gestire in modo opportuno. Di grande aiuto risulta, a tal fine, il poter disporre di dati statistici, in particolare sui valori e gli andamenti della produttività e sui tassi di errore riscontrati.

In questo articolo viene fatta una sintesi dei rilevamenti sperimentali effettuati esaminando un numero significativo di casi; i dati sono analizzati per determinare se, e quali, correlazioni esistono tra i principali fattori in gioco.

2. Misure di produttività

Alcuni risultati sperimentali sulla produttività provengono dalla raccolta di dati fornita dal Rome Air Development Center (RADC) del System Command Center dell'Aeronautica degli Stati Uniti. RADC ha raccolto ed analizzato (Nelson, 1978) dati di 407 progetti per un totale di più di 21 milioni di DSLOC (Delivered Source Lines of Code: linee di codice sorgente consegnate), con una dimensione media dei progetti di circa 52K DSLOC, comportanti l'uso di 28 linguaggi diversi per il loro sviluppo. I dati raccolti si riferiscono ad una grande varietà di problemi (elettronica aerospaziale, calcolo numerico, elaborazione di dati commerciali, sistemi di comando e controllo, sistemi di guida di missili), studiati da gruppi diversi di fornitori U.S.A. per il Ministero della Difesa Americana e sono di sviluppo recente (fine anni settanta). I dati vennero analizzati e, ove possibile, i risultati relativi al numero di errori e al livello di produttività vennero raggruppati per linguaggio e tecnica di sviluppo usati. Il valore di produttività medio, calcolato su 379 progetti per cui i dati di produttività sono disponibili, è di 388 linee di codice sorgente per mese uomo (DSLOC/MM); il numero di errori misurati per i progetti medesimi verrà discusso più avanti.

2.1. Produttività e dimensioni dei programmi

Un tentativo di correlazione tra produttività e dimensioni del progetto, fatto da Nelson (1978), dimostra che una tale dipendenza non esiste. Infatti l'interpolazione con il metodo dei minimi quadrati mostra una retta quasi parallela alle ascisse dove il fattore di correlazione è .0334, troppo basso per implicare qualunque dipendenza. L'unica conclusione che

Fig. 1 - Dati di produttività in funzione del linguaggio di programmazione usato (Bibl. [9]).

LINGUAGGIO	NUMERO DI PROGETTI	DIMENSIONE MEDIA DEL PROGETTO (DSLOC)	PRODUTTIVITÀ MEDIA (DSLOC/MM)	DIMENSIONE TOTALE DEL CAMPIONE
ASSY	244	30.7K	420	7500K
FORTRAN	25	10.7K	321	270K
COBOL	33	27.1K	566	890K
PL/I	22	40.7K	391	890K

risulta dall'analisi di Nelson è che più grande è il progetto, più persone verranno impiegate e più a lungo.

2.2. Produttività e linguaggio di programmazione

L'effetto che il linguaggio di programmazione ha sulla produttività, secondo i dati forniti da Nelson (1978), è esposto in fig. 1. È interessante notare che i dati presentati non permettono nessuna conclusione logica; l'uso del COBOL sembra dare un aumento di produttività, ma l'uso del FORTRAN e del PL/1 dà una riduzione rispetto al linguaggio assemblatore (ASSY). È certo possibile che il tipo di problema affrontato da gruppi di linguaggio diverso abbia contribuito in modo determinante; così, per esempio, ASSY e PL/1 possono essere stati adoperati per lo sviluppo di sistemi di difficoltà molto maggiore di quella dei problemi affrontati in COBOL. Per di più, non è facile da stabilire se il numero di DSLOC che si servono di un linguaggio di alto livello per conseguire una funzionalità specifica, è molto inferiore del corrispondente numero di DSLOC in linguaggio assemblatore. Un ulteriore elemento di confusione può essere il fatto che il termine linguaggio assemblatore in realtà si riferisce ad una gran varietà di linguaggi, diversi sia per la diversa architettura dello hardware sia per il più o meno ricco repertorio di macro istruzioni.

Malgrado, queste incertezze, le differenze riscontrate tra i vari linguaggi non significano che un particolare linguaggio sia migliore di un altro dal punto di vista della produttività del processo di sviluppo.

Risultati simili sulla relativa indipendenza dello sviluppo software dal linguaggio usato, sono stati presentati da Boehm (1980), che ha studiato i risultati di un progetto sperimentale in cui gruppi diversi hanno usato linguaggi diversi.

2.3. Produttività e tecniche di sviluppo

L'effetto della tecnica di sviluppo usata sulla produttività è dato in Fig. 2. I numeri tra parentesi indicano il numero di progetti esaminati in ciascuna categoria, il termine PT significa programmazione tradizionale ed indica progetti sviluppati senza l'aiuto di quelle che vengono considerate discipline di programmazione moderna (PM), come diagrammi HIPO, strutture di controllo, sviluppo top-down, ecc. Di 379 progetti totali, 210 sono classificati PT.

I risultati contrassegnati PM riguardano progetti che, per almeno il 90% dei casi, sono stati sviluppati con l'uso di discipline di programmazione moderna. La figura 2 mostra che 58 progetti, su un totale di 379, sono classificati PM. Il termine Specs indica l'uso di un linguaggio formale per la descrizione del progetto.

Bisogna notare che tutte le discipine di programmazione moderna mettono l'accento sulla fase di progetto e cercano di guidare il processo di programmazione facendo uso di risultati ottenuti dalla ricerca teorica. L'uso di un linguaggio formale (o semiformale) per esprimere le decisioni del progetto, può essere interpretato come un tentativo di aumentare l'importanza delle decisioni di progetto rispetto all'intero processo di sviluppo. In altre parole, la PM cerca di introdurre metodi e discipline analoghe alla moderna ingegneria in cui l'enfasi dello sviluppo risiede nella fase di definizione.

La figura 2 fornisce la media e la mediana dei dati analizzati. Data la forma asimmetrica della curva statistica, la mediana rappresenta un dato migliore della media. Inoltre, il test

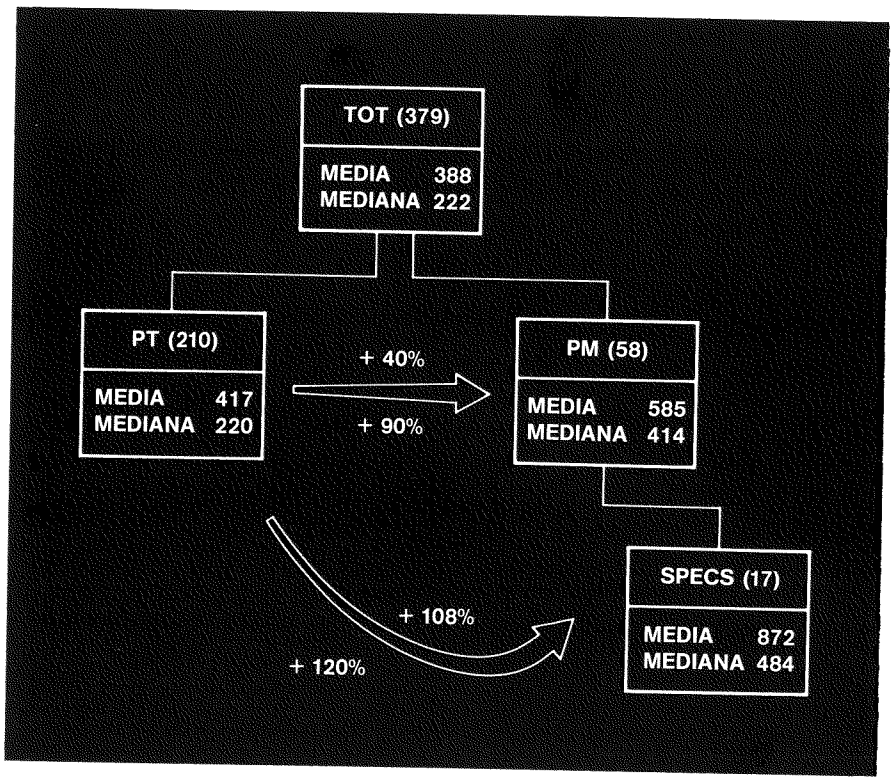


Fig. 2 - Effetto delle tecniche di sviluppo (Bibl. [9]).

per errori di tipo II dimostra che la probabilità di errore, accettando i miglioramenti dichiarati per PM e Specs, è del 18% e 26% rispettivamente, con un livello di confidenza di 0.05.

Il confronto tra PT e PM mostra un aumento di produttività del 90% sulla mediana e del 40% sulla media. Più imponente è l'aumento mostrato da Specs su PT: 108% e 120% per media e mediana, rispettivamente.

2.4. Conclusioni

Per concludere, i dati esaminati presentano le seguenti caratteristiche:

1. È impossibile stabilire una correlazione tra dimensione del prodotto e produttività sulla base dei dati forniti dai tipi di progetto analizzati da Nelson (1978) (cioè sviluppi di relativamente grandi dimensioni)

2. L'adozione di linguaggi diversi non dà risultati di chiara efficacia per la produttività nello sviluppo di grandi sistemi

3. Le moderne tecniche e discipline di programmazione rappresentano una via interessante per un aumento di produttività.

4. Tra le varie tecniche PM, la rappresentazione del progetto in un linguaggio formale (Specs) sembra la più efficace per un aumento di produttività.

La prima conclusione non risulta intuitiva. Essa sembra suggerire che la produttività sia una caratteristica dell'organizzazione di sviluppo piuttosto che del prodotto. Le conclusioni 2 e 3 sono evidenti per chi ha esperienza nello sviluppo di grandi sistemi. La conclusione 4 è piuttosto sorprendente nel senso che, ad un esame superficiale, una descrizione formale del progetto sembrerebbe far aumentare il lavoro, anzichè diminuirlo. Il suo uso conduce invece ad un raddoppio di produttività.

L'unica conclusione possibile che si può trarre da questi risultati è che una documentazione non equivoca degli elementi del progetto durante le fasi iniziali di sviluppo, ha un impatto favorevole sulla produttività. L'uso di un linguaggio di programmazione appropriata aiuterà certamente l'implementazione del sistema ma, non avendo il linguaggio un ruolo importante durante la fase di progetto, ciò avrà poco effetto sulla produttività totale.

3. Esperimenti sul tasso di errori

La raccolta di dati sperimentali da parte del RADC (Nelson, 1978), fornisce ulteriori dati sulla percentuale di errori, dove gli errori sono definiti come difetti segnalati ufficialmente, prima della consegna (release) del software agli utilizzatori.

3.1. Tasso di errori e dimensioni del progetto.

I dati del RADC sul tasso di errori, misurato come il numero di errori per mille DSLOC, dimostra la stessa mancanza di correlazione con le dimensioni del progetto già vista per la produttività. Anche in questo caso l'interpolazione con l'uso dei minimi quadrati, dà una curva quasi piatta, (leggermente discendente con l'aumento delle dimensioni del progetto), con un coefficiente di correlazione così basso da scoraggiare qualunque conclusione. (Ovviamente il numero totale di errori risulta funzione delle dimensioni del progetto). Ciò sembra, di nuovo, suggerire che il tasso di errori dipende dalla organizzazione di sviluppo (o processo di sviluppo) piuttosto che dalle caratteristiche del prodotto. Sembrerebbe che il processo possa gestire un determinato tasso di errori (così come un certo livello di produttività) e che quindi un maggior tempo di sviluppo sia necessario per gestire il maggior numero totale di errori presentati dai grandi progetti.

3.2. Tasso di errori e linguaggio di programmazione

La fig. 3 presenta dati di errore in funzione del linguaggio. Questi dati possono ingannare; infatti i valori medi sono esattamente ciò che dice il termine, cioè i valori medi sono validi soltanto in media. Qualora si presentino grandi varianze, i valori medi non forniscono nessuna informazione utile. Si può dire, tuttavia, con sicurezza, che non c'è alcuna evidenza che un particolare linguaggio abbia un effetto decisivo sul tasso di errori. Alcuni linguaggi, come FORTRAN e COBOL, hanno un

Fig. 3 - Dati sul tasso di errore in funzione del linguaggio usato. (Bibl. [9]).

LINGUAGGIO	ERRORE MEDIO	DIMENSIONE DEL CAMPIONE (N. DI LAVORI)
ASSY	26.6	244
FORTRAN	15.1	25
COBOL	12.9	33
PL/I	33.3	22
JOVIAL J4	16.4	2
JOVIAL J3	50.3	4
JOVIAL J3TS	17.0	1
CENTRAN	19.4	11

tasso di errori inferiore a quello dell'assemblatore; ma PL/1 e JOVIAL L3 si comportano in modo opposto.

Se i numeri disponibili per tutti i linguaggi di alto livello sono messi assieme, la media di errori è di 23,5 per mille linee di codice (dimensione media del programma: 37.9K linee). Questo valore medio rappresenta uno scarso miglioramento rispetto agli assemblatori (26.6 errori per mille linee, con una dimensione media del lavoro di 30K) ed è impossibile concludere che i linguaggi di alto livello (HLL) rappresentino un fattore significativo nella riduzione del tasso di errori.

È tuttavia, opinione comune che l'uso di HLL riduca il tasso di errori. In effetti, non c'è dubbio che questo si verifica nel caso dello sviluppo di programmi individuali. L'indicazione fornita da questi dati è, di nuovo, che la programmazione di grandi sistemi software è influenzata da fattori che non sono sostanzialmente influenzati dalle tecniche di codifica, ma sono caratteristici del processo di sviluppo, operazione questa che include un notevole lavoro per l'integrazione dei vari moduli nel sistema. È esperienza comune che molti errori sono scoperti solo in questo ultimo stadio dello sviluppo.

Hamilton e Zelding (1976), in un altro esperimento, misurarono la proporzione di errori scoperti nelle fasi di integrazione e di prova del sistema. Essi trovarono che il 73% dei problemi riscontrati nel software del progetto Apollo erano collegati alla fase d'integrazione. Conclusioni simili, con riguardo a grandi sistemi operativi sono riportate da Richte (1979).

Questi esperimenti non significano che gli HLL non siano di nessuna utilità nello sviluppo di sistemi software di grandi dimensioni, ma stanno ad indicare che lo sviluppo di grandi sistemi con una bassa percentuale di errori richiede qualcosa di più sofisticato degli HLL disponibili oggi. Ed inoltre, non è possibile trarre alcuna conclusione sulla possibilità di creare nuovi HLL che risultino efficaci per la riduzione del tasso di errori.

3.3. Tasso di errori e tecniche di sviluppo

I dati raccolti da RADC non analizzano l'effetto delle tecniche di programmazione e/o dell'uso di un linguaggio di progetto sulla percentuale di errori. Nello studio di Vessey e Weber (1983), limitato ad un gruppo di programmi in COBOL scritti da tre organizzazioni diverse, l'effetto di metodi moderni di programmazione sul tasso di errori sembra trascurabile. Non ci sono dati disponibili riguardanti l'uso di un linguaggio di progetto.

Un'altra serie di osservazioni sperimentali sulla percentuale di errori chiarifica ulteriormente il processo di sviluppo software. Jones (1979) presenta una serie di dati che puntualizzano i problemi incontrati nello sviluppo dei grandi sistemi. Come dall'esame della figura 4, il diagramma tracciato, visto come andamento piuttosto che in termini di valori numerici, indica che, tra le molte cause che contribuiscono alla produzione di errori, l'incremento di errori dovuto a specifiche di disegno logico, in relazione alle dimensioni del sistema, è chiaramente il più significativo. Jones mette questo risultato in relazione con un altro dato sperimentale: la percentuale di funzionalità interne di un sistema (cioè le decisioni di progetto) documentate *in un modo qualsiasi,* diminuisce con le dimensioni del sistema. C'è da notare, inoltre, che l'aumento del tasso di errori è più rapido per sistemi di dimensioni da 1K a 50K linee, che può essere considerata la linea di demarcazione tra sistemi di software grandi e piccoli.

Altri dati sperimentali mostrano che la tendenza verso l'errore è propria di relativamente pochi moduli di un sistema software.

Sia Jones (1983) che Zaharchuck (1980) rilevarono che la maggior parte dei moduli presentano un livello di errori vicino a zero durante la fase di prova del sistema. (Jones analizzò molti sistemi software di tipo diverso; Zaharchuck, invece, seguì parecchie release di uno stesso sistema operativo). Jones dichiara che il 60% dei moduli di un grande sistema sono, con buona approssimazione, privi di errori e fornisce la regola pratica che il 50% degli errori di un sistema deriva dal 5% dei moduli del sistema stesso. I dati forniti da Zaharchuck, inoltre, indicano che i moduli più suscettibili di errore sono gli stessi nelle successive release, anche quando esse presentino l'aggiunta di funzionalità diverse.

Questi moduli "proni all'errore" non possono essere associati con nessuna particolare tecnica di sviluppo o linguaggio di programmazione, dato che spesso sono codificati, nello stesso sistema, con linguaggi diversi. Inoltre le loro dimensioni non si scostano sostanzialmente da quella media degli altri moduli del sistema. La loro suscettibilità all'errore sembra essere una caratteristica specificamente connessa al loro ruolo nel sistema.

4. Comunicazione delle decisioni di progetto

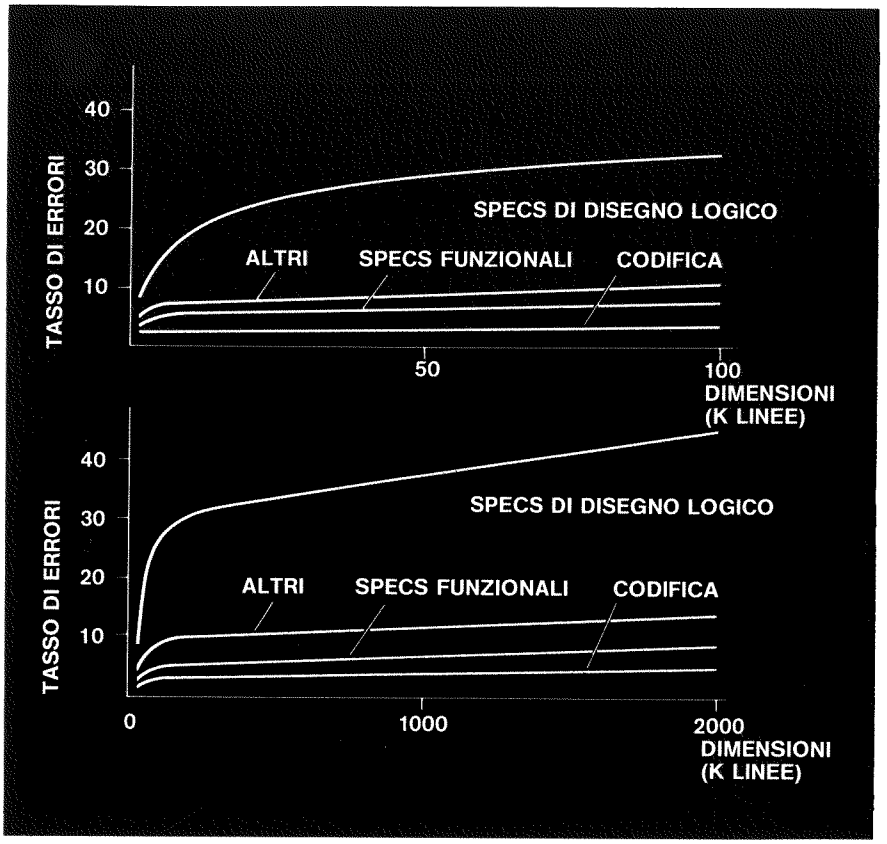
La descrizione delle funzioni che un sistema software deve realizzare è un aspetto molto importante della fase di progetto, in cui le specifiche del sistema software devono essere definite e scritte. Sfortunatamente, anche quando le specifiche sono scritte, [e per la maggior parte dei casi nei grandi sistemi non lo sono, secondo Jones (1979)] non sono umanamente leggibili, date le dimensioni di un tale documento che, anche se propriamente organizzato, richiederebbe una vita per essere letto ed interpretato. Tale documento scritto, avente come scopo lo scambio di informazioni tra progettisti, è un compito impossibile per ovvie ragioni organizzative.

Ne risulta che il modo più comune di comunicazione per lo scambio di decisioni riguardanti il progetto, è quello verbale.

Il problema delle comunicazioni verbali risiede nella imprecisione intrinseca del linguaggio umano ed è verosimile che incomprensioni derivanti da tali comunicazioni circa le interazioni dei moduli, siano una causa di errori.

Una soluzione molto nota per questo problema è il "chief programmer team" (Baker, 1972), con cui si cerca

Fig. 4 - Andamento del tasso di errore in funzione delle dimensioni del programma. (Bibl. [6]).



di potenziare le capacità di un programmatore liberandolo da impegni che non siano direttamente collegati alla creazione del sistema. Il codice viene scritto dai vari programmatori, mentre la globalità del compito è presente solo nella mente del "chief programmer". Il caso di un evento catastrofico (il "chief programmer" passa ad un'altra società, un'epidemia d'influenza, etc.) viene aggirato fornendo le stesse informazioni ad un programmatore di ricambio che possa assumere funzioni direttive nello sviluppo, se necessario.

L'esperienza insegna che questo schema è ingegnoso ed efficace, però con una limitazione. L'efficacia deriva dal fatto che una tecnica ben sperimentata (l'abilità, cioè, di un programmatore solitario di realizzare correttamente ed efficientemente un prodotto) viene estesa ad un campo più ampio di applicazione, semplicemente aumentando l'efficenza lavorativa. La limitazione consiste nel fatto che questa idea è buona nella misura in cui lo è la mente del chief programmer, e la sua capacità di dirigere, comunicare e controllare altri progettisti. Inoltre, se il sistema è troppo vasto per le capacità di una sola persona, l'idea diventa irrealizzabile. È stata anche proposta l'idea di avere dei sottogruppi, cioè una gerarchia di chief programmers, che non sembra, però, possibile realizzare (l'applicazione non è stata riportata nella letteratura) a causa delle enormi difficoltà di comunicazione.

La discussione svolta finora suggerisce l'interpretazione che l'uso di un linguaggio formale per esprimere le decisioni del progetto, è sufficiente a fornire la concisione e la precisione necessarie. In effetti, i dati riguardanti l'aumento di produttività in sviluppi che usano un linguaggio formale per il progetto sembrano sostenere quest'ipotesi. Tuttavia, sarebbe utile che ogni linguaggio formale usato per descrivere le decisioni del progetto, potesse essere usato anche per pronosticare la correttezza del progetto stesso. L'informatica teorica si è occupata del problema della verifica della correttezza di un programma senza doverlo eseguire. Le tecniche studiate possono essere introdotte nel processo di sviluppo, aumentando la potenzialità previsionale in fase di progetto.

5. Esperimenti sul ciclo di vita del software

Il lavoro sperimentale più chiarificatore sul ciclo di vita dei grandi sistemi software è stato fatto da Lehman e Belady (1979), che dimostrarono come tali sistemi siano, durante il loro uso, invariabilmente soggetti a modifiche dovute a cambiamenti continui del contesto (requisiti dell'utente, modifiche di hardware e/o software, etc.) e che la quantità di lavoro di modifica eseguito sul sistema per unità di tempo *è essenzialmente costante,* qualora mediato su lunghi periodi di tempo. Questa ricerca ha dimostrato sperimentalmente una precisa correlazione lineare tra la quantità totale di modifiche e migliorie apportate al sistema e l'età del sistema stesso. Lo studio ha, inoltre, provato che qualora la curva rappresentante il lavoro cumulativo sul sistema, venisse forzata (presumibilmente in seguito ad interventi del management) verso l'alto per un certo periodo di tempo, l'aumento era poi seguito da un calo, per cui restava costante il valore medio del lavoro eseguito per unità di tempo.

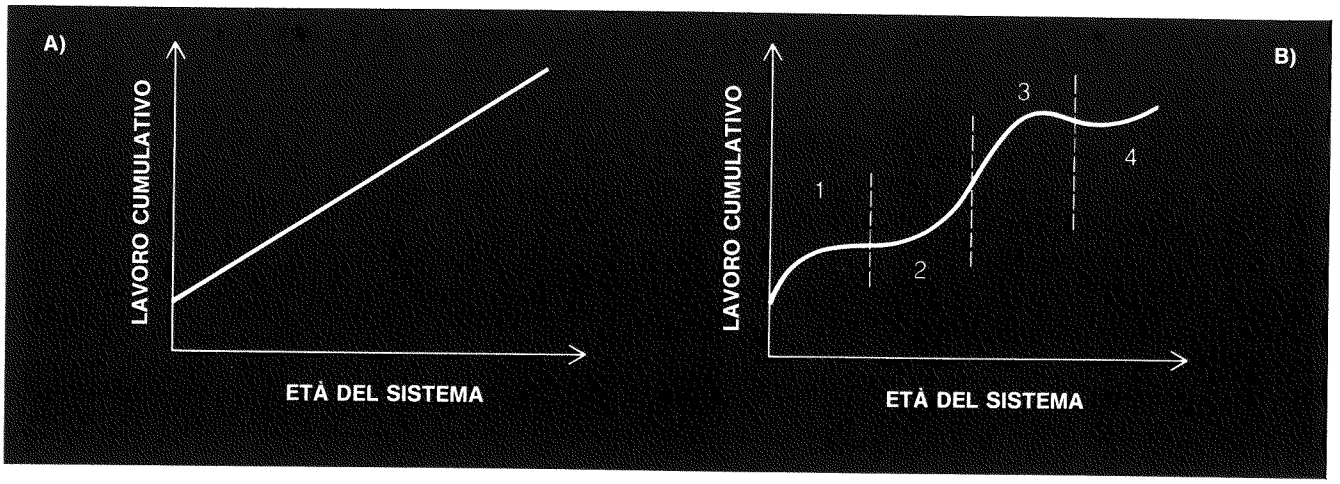


Fig. 5 - Dinamica dei grandi sistemi di software: (A) lavoro totale effettuato in rapporto all'età del sistema; (B) effetti di secondo ordine.

C'è da notare che tale valore medio non è influenzato da variazioni della quantità delle risorse usate, delle tecniche di sviluppo, del linguaggio di codifica etc. In realtà, un grande sistema possiede una dinamica intrinseca che è influenzata solo in modo limitato e di breve durata da fattori esterni (il personale impiegato, per esempio). Per cui risultano molto frustranti per i manager del software le imposizioni di accelerare i tempi.

L'evidenza di un costante lavoro di modifica può essere interpretata in accordo con un'opinione espressa precedentemente e cioè che la mancanza di correlazione tra produttività e tasso di errori rispetto alle dimensioni del progetto sta ad indicare che questi due parametri sono dipendenti dalle caratteristiche del processo di sviluppo piuttosto che da quelle del prodotto, quali le dimensioni o la tecnica di codifica.

C'è tuttavia da sottolineare che il lavoro di modifica dipende, a sua volta, da caratteristiche del prodotto stesso, come comprensibilità del codice e organizzazione del sistema.

Lo studio di Lehman e Belady ha messo in evidenza un'altra caratteristica dei grandi sistemi: il processo di continua modifica comporta un deterioramento della struttura concettuale del sistema, rendendo il lavoro da eseguire sempre più arduo, il codice esistente più difficile da comprendere ed aprendo la strada all'introduzione di nuovi errori dovuti ai cambiamenti e alla correzione di altri errori. Per questo motivo la produttività del processo decresce col tempo.

Per controbilanciare questo deterioramento strutturale, diventa necessario ricostruire la struttura concettuale del sistema. Questo lavoro di ristrutturazione aumenta però il costo del sistema, considerato sul suo ciclo di vita, e non può essere eseguito ad una velocità superiore a quella permessa dalla dinamica intrinseca del sistema stesso, il che a sua volta ritarda l'introduzione di modifiche.

I concetti espressi in precedenza possono essere meglio compresi con l'esame dei dati riportati nella figura 5. La figura 5A mostra un tipico grafico del lavoro totale fatto, in funzione della età del sistema. La relazione lineare sta a significare che il lavoro eseguito per unità di tempo è costante. La Figura 5B mostra un effetto di second'ordine, trascurato nella Figura 5A: durante i periodi 1 e 3, forse a seguito di decisioni del management, viene fatto uno sforzo per effettuare più lavoro per unità di tempo. (La curva mostra una pendenza maggiore di quella accettata spontaneamente dal sistema). Invariabilmente, periodi del tipo 1 e 3 sono seguiti da altri del tipo 2 e 4, dove la curva si appiattisce (meno lavoro viene eseguito per unità di tempo) per compensare l'alta produzione dei cicli precedenti e, presumibilmente, per eseguire il lavoro di ristrutturazione necessario per accettare nuove modifiche e migliorie.

Tuttavia, in alcuni casi, il management può decidere di non impiegare del personale per il lavoro di ristrutturazione, con la conseguenza di un calo di produttività per gli operatori del sistema; e quindi l'impossibilità di aggiungere nuove funzionalità.

Data l'assenza di una documentazione completa e formale del progetto, gli sviluppatori, per eseguire modifiche sul sistema, dovranno leggere il codice per poter stabilire con precisione la logica del progetto. È ragionevole assumere che il lavoro di manutenzione sia, in qualche modo, influenzato dal linguaggio di programmazione. Studi sperimentali (Boehm, 1981; Lientz e Swanson, 1981) mostrano, invece, che il numero di istruzioni di codice sorgente modificato per uomo-mese è essenzialmente indipendente dal linguaggio di programmazione usato. La difficoltà di manutenzione di un grande sistema di software è essenzialmente funzione della sua organizzazione. Sistemi diversi, sviluppati con lo stesso linguaggio e persino con lo stesso hardware, danno costi di manutenzione diversi. Perciò, nei termini forniti dalle osservazioni di Lehman e Belady, si

può dire che l'inclinazione della curva nella figura 5A è indipendente dal linguaggio o dall'hardware usati per lo sviluppo del sistema. Quest'affermazione è valida solo per i linguaggi correntemente in uso; non c'è modo di stabilire, per il momento, la possibilità di creare un linguaggio di programmazione che riesca ad influenzare la pendenza della curva di figura 5A.

Questi dati sperimentali rivelano un'altra caratteristica dei grandi sistemi, di non facile intuizione: essi si comportano come canali di portata fissa; quando la portata massima è raggiunta, niente può essere più aggiunto; per di più questa portata si riduce col tempo, a meno che del lavoro specifico non venga eseguito per mantenerla al livello originale.

Sistemi diversi mostrano valori diversi di capacità ed alcuni accettano più lavoro per unità di tempo di altri, per cui si deve assumere che l'organizzazione architettonica del sistema fissi la sua capacità di accettare modifiche e migliorie.

6. Conclusioni

I risultati sperimentali dimostrano che la produttività non è sostanzialmente influenzata dalle dimensioni del progetto o da strumenti di sviluppo come HLL. Le più recenti tecniche di programmazione, tuttavia, che danno rilievo alla fase di progetto, sembrano realizzare un aumento di produttività. Anche la percentuale di errori è sostanzialmente indipendente dalle dimensioni e dal linguaggio del progetto.

Questa invarianza della produttività e del tasso di errori può essere interpretata come l'indice di una rigidità intrinseca del processo di sviluppo dei sistemi di software. La difficoltà di comunicazione tra progettisti può contribuire a tale rigidità.

Un'altra importante osservazione sperimentale è che un numero limitato di moduli è responsabile della maggior parte degli errori di un sistema; questo dimostra che alcune caratteristiche del ruolo di tali moduli comportano un impatto decisivo sulla affidabilità complessiva.

Il comportamento dei sistemi software durante il loro ciclo di vita è stato analizzato sperimentalmente; i risultati hanno ulteriormente delineato alcune caratteristiche di rigidità del processo e cioè che un sistema accetta solo una quantità fissa e costante di modifiche e che tali modifiche comportano un deterioramento nella sua struttura concettuale.

7. Bibliografia

[1] BAKER, F.T. 1972, Chief programmer team management of production programming. *IBM System Journal,* Vol. 11, no. 1 : 56-73.

[2] BELADY, L. A., and LEHMAN, M. M. 1979. The characteristics of large systems. In *Research directions in software technology,* edited by P. Wegner. Cambridge, Mass.: MIT Press.

[3] BOEHM, B. W. 1980. Developing small-scale application software products. In *Proceedings of the 1980 IFIP Congress,* edited by S. H. Lavington, pp. 321-326. North-Holland Publishing.

[4] BOEHM, B.W. 1981. *Software engineering economics.* Englewood Cliffs, N.J.: Prentice-Hall.

[5] HAMILTON, M., and ZELDING, S. 1976. Higher order software - a methodology for defining software. *IEEE Transactions on Software Engineering* SE-2 (no. 1): 9-32.

[6] JONES, C. 1979. A survey of programming and specification techniques. In *Proceedings of the Conference on Specification of Reliable Software, Cambridge, Mass.,* pp. 91-103.

[7] JONES, C. 1983. An ITT internal presentation as reported by E. Kirby at an RASA-SU meeting, Phoenix, Ariz.

[8] LIENTZ, B. P., and SWANSON, E. B. 1981. Problems in application software maintenance. *Communications of ACM 24* (no. 11): 763-769.

[9] NELSON, R. 1978. Software data collection and analysis. Draft of a partial report, Rome Air Development Center, Air Force Systems Command, Griffiss Air Force Base, N.Y.

[10] RICHTER, J.C. 1979. Note on testing effectiveness. Honeywell Information Systems internal note.

[11] VESSEY, G., and WEBER, R. 1983. Some factors affecting program repair maintenance: An empirical study. *Communications of ACM 26* (no. 2); 128-134.

[12] ZAHARCHUCK, A. 1980. Error-prone modules, Honeywell Informations Systems internal memo.

NOTA: Questo articolo riproduce un capitolo del libro dell'Autore: "Development and maintenance of large software systems", ed. Wadsworth, 1984. La versione italiana del volume verrà pubblicata nella collana di testi "Quaderni di Informatica" (F. Angeli ed.).

Una metodologia di misura della qualità del software

MARGHERITA KARLIN

Honeywell Information Systems Italia
Centro di Ricerca e Progettazione
Pregnana Milanese

1. Introduzione

Il giudizio di qualità su di un prodotto software, al momento del suo rilascio, è sempre stato un problema di non facile soluzione per le case di produzione in quanto prevalentemente arbitrario e soggettivo per coloro che, come responsabili nell'organizzazione, devono esprimere tale giudizio.

Il motivo principale di tale arbitrarietà è che storicamente si è guardato al software come a qualcosa di intangibile, di indefinito e quindi in sostanza, di *non misurabile;* la mancanza di un *metro di misura* univoco e chiaro ha fatto si che la valutazione fosse non rigorosa ma soggettiva.

Quando anche si è cominciato a definire una metodologia di testing precisa, come quella in uso presso la nostra ingegneria software, basata su check-list, sia topologiche che funzionali,la grande, per non dire infinita, varietà di casi teoricamente da provare unita alla frequente variabilità dell'ambiente stesso di prova nelle fasi finali prima del rilascio, ha fatto sì che anche un approccio metodologico basato su elementi misurabili (numero di elementi nella check-list, numero di test ecc. ec.) fosse di fatto realisticamente insufficiente a fornire *misure di qualità* oggettivamente valide e applicabili a tutti i tipi di prodotti.

Ciò che si è fatto, allora, è di tradurre la Metodologia di Testing in una Metodologia di Misura della Qualità che, derivata da quella, consentisse però di effettuare delle misurazioni fattibili e oggettive.

È stata così definita una *convenzione* di Copertura del Controllo di Qualità basata su elementi misurabili.

Questa consiste nel:

- percorrere almeno una volta tutti i rami del codice
- esercitare almeno una volta tutti i parametri visibili all'utente
- emettere almeno una volta tutti i messaggi

Il primo punto costituisce la "copertura topologica minima", mentre gli altri due la "copertura funzionale minima".

Tale convenzione deve essere ripetuta nell'ambiente che un Comitato di Controllo Modifiche indica come quello "finale". Inoltre per modifiche dovute a errori rilevati in questa fase, è richiesto che vengano ripetuti almeno

- i test che sollecitano i rami modificati
- tutti i test relativi alla visibilità utente
- i test relativi ai messaggi eventualmente coinvolti.

2. Indice di difettosità interna

Si è quindi definito un *indice di difettosità interna,* basato sul numero di difetti rilevati sul prodotto prima del rilascio e sulla loro gravità e dipendente anche, in misura diversa, dalla fase di produzione (analisi, codifica, qualificazione) in cui il difetto è stato rilevato.

Tale indice dipende anche in modo inversamente proporzionale dalla "quantità di codice" che costituisce il prodotto e da un indice di sollecitazione che in pratica indica quanto la *convenzione* è stata applicata.

L'Indice di Difettosità Interna può quindi, anche intuitivamente, essere utilizzato ai fini della misura della qualità in quanto ovviamente a maggiore difettosità corrisponde minore qualità, e viceversa.

Per come è stato definito, esso ha il vantaggio di poter essere calcolato.

Infatti ha:

$$D_I = D/IS_I$$

ove

D = *valore di difettosità totale interna*

IS_I = *indice di sollecitazione interna*

La Difettosità Totale Interna D viene calcolata conteggiando il numero di difetti rilevati sul prodotto e assegnando ad essi un peso diverso a seconda della fase del ciclo di produzione in cui si sono scoperti, ossia secondo la seguente formula:

$$D = \sum_{I=1}^{3} F_I\,[10S_I + 3M_I + T_I]/KSLOC$$

Dove:

I = *numero progressivo della fase*

S_I = *numero totale di difetti "gravi" trovati nella fase I*

M_I = *numero totale di difetti "medi" trovati nella fase I*

T_I = *numero totale di difetti "trascurabili" trovati nella fase I*

KSLOC = *migliaia di linee codice sorgente del prodotto/release*

F_I = *valore costante relativo alla fase I*

in cui le fasi 1, 2, 3 e il relativo valore F_I sono:

1) *Analisi/codifica/debugging* $F_1 = 1$

2) *Integrazione/quality control* $F_2 = 7$

3) *Quality assurance* $F_3 = 50$

L'indice di Sollecitazione Interna IS_I viene calcolato in base alla seguente definizione:

$$IS_I = (R + P + M)/3$$

essendo:

R = *% di rami percorsi dal testing nella qualificazione interna*

P = *% di parametri esterni sollecitati dal testing nella qualificazione interna*

M = *% di messaggi emessi nel testing nella qualificazione interna*

È chiaro che

$$0 < IS_I < 1$$

e anche intuitivamente si capisce quindi che se la sollecitazione è massima (= 1) la Difettosità Interna è uguale a quella Totale e, a parità di difetti rilevati, ha il valore minimo, mentre se la sollecitazione è minima (tendente a zero) la Difettosità Interna o non è significativa, o, a parità di difetti rilevati,tende ad aumentare all'infinito.

C'è da aggiungere che il giudizio circa la gravità del difetto rilevato, ossia *grave, medio e trascurabile,* viene assegnato, in modo indubbiamente obiettivo, sia per i difetti rilevati durante la qualificazione interna che, come vedremo, per quelli rilevati in field, solamente in base alle conseguenze del difetto, ossia:

- dati utente rovinati: *grave*
- sistema fermo che necessita ri-inizializzazione: *medio*
- ogni altro caso: *trascurabile*

Essendo tutti i dati che ricorrono nelle formule misurabili, è chiaro che si può avere, per ogni prodotto, un valore che indichi il suo *indice di difettosità.*

Questo da solo non ci permette ancora un giudizio di qualità, anche se intuitivamente lo può rappresentare, se non si fa un passo ulteriore che è quello di confrontare tale indice relativamente a prodotti diversi, e verificare la coerenza del loro comportamento dopo il rilascio (quindi a posteriori) rispetto ai loro indici calcolati prima del rilascio.

Ciò ci consentirà di dare una validazione al metodo, o di affinarlo a fronte dell'esperienza.

3. Indice di difettosità esterna

Per verificare il comportamento dei prodotti dopo il rilascio, si applica lo stesso approccio usato per la Qualificazione interna, dando però un diverso (maggiore) peso alla gravità dei difetti rilevati, e assumendo come *indice di sollecitazione* semplicemente il *numero totale di utenti* reali di quel prodotto.

In pratica *l'indice di difettosità esterna* D_F è definito, in modo analogo all'altro:

$$D_F = D/IS_F$$

dove D *è la difettosità totale in field,* data da:

$$D = F_F (10\ S + 3\ M + T)/KSLOC$$

in cui:

S = *numero totale progressivo di difetti "gravi" rilevati in field*

M = *numero totale progressivo di difetti "medi" rilevati in field*

T = *numero totale progressivo di difetti "trascurabili" rilevati in field*

KSLOC = *migliaia di linee codice sorgente del prodotto/release*

F_F = *100 assunto come costante analoga al valore del costo relativo di correzione in field*

IS_F è *l'Indice di sollecitazione in field,* ed è *assunto uguale al numero di sistemi installati in field, che utilizzano il prodotto/release.*

4. Primi risultati della sperimentazione

Nell'Ingegneria Software della HISI si è cominciato ad applicare questa metodologia di Misura della Qualità del software da circa due anni, sviluppando anche strumenti che consentissero e di registrare tutti i dati relativi alla problematica in oggetto e di calcolare le conseguenti misure sui prodotti sviluppati e/o già rilasciati.

I primi risultati tendono a dimostrare che il comportamento di un prodotto dopo il rilascio non è molto dissimile da quello prima del rilascio, il che ci porterebbe a dedurre che se un prodotto ha un alto Indice di Difettosità Interno, lo avrà tale anche in Field e quindi non è di "buona qualità".

Portando ad esempio due prodotti emessi dalla HISI rispettivamente nel giugno '84 e giugno '85 (di cui il secondo "B" alla sua uscita ha praticamente sostituito in field il primo "A") si nota dalle misure fatte secondo quanto sopra esposto:

D_I (A) = 23.07

D_I (B) = 2.45

con un rapporto di circa 9 a 1, mentre per la misura della Qualità in field:

D_F (A) = 0.29 (dopo ~ 1 anno)

D_F (B) = 0.04 (dopo ~ 8 mesi)

con un rapporto di circa 7 a 1 che, non essendo il prodotto B ancora a regime e comunque essendo rilasciato da meno di un anno, potrebbe nei prossimi mesi, secondo la tendenza già mostrata, avvicinarsi al rapporto 9 a 1 tra gli Indici di Difettosità Interna se l'Indice di Difettosità Esterna di B continuerà a diminuire.

Con ciò si dimostrerebbe che il rapporto tra le Difettosità di due prodotti prima del rilascio si mantiene proporzionale anche dopo la loro installazione in field.

Dalle misure effettuate, si nota anche un altro aspetto molto interessante che, se ulteriormente confermato in futuro, potrà davvero porta-

re utili indicazioni previsionali sul comportamento di un prodotto in field.

Più precisamente, si può notare che anche su prodotti diversi il rapporto tra Difettosità Interna e Difettosità Esterna, (almeno per prodotti del tipo e delle dimensioni dei suddetti "A" e "B") tende ad essere molto simile, ossia:

D_I (A) = 23.07 D_F (A) = 0.29

D_I (A)/D_F (A) = 79

D_I (B) = 2.45 D_F (B) = 0.04

D_I (B)/D_F (B) = 61

ove, come detto sopra, se D_F (B) diminuirà, il rapporto $D_I(B)/D_F(B)$ crescerà, e quindi si avvicinerà al valore 79.

Se le future misurazioni, su questi ed altri prodotti, confermeranno quanto sembra di poter ipotizzare dalle misure fatte, si arriverebbe, sperimentalmente, ad ottenere un valore K come costante del rapporto:

D_I/D_F

per un qualsiasi prodotto del tipo e delle dimensioni di A e B.

Avendo tale valore, si potrebbe dedurre per un prodotto non ancora rilasciato avente Difettosità Interna D_I, il valore previsionale della sua Difettosità Esterna:

$D_F = K/D_I$

e quindi, conoscendo, assumendo o estrapolando il numero di utenti, si otterrebbe una previsione del numero di difetti che saranno rilevati su tale parco, da cui si potrà giudicare a priori la Qualità del prodotto in field.

Ciò non può che essere di aiuto a indicare e promuovere tempestivamente azioni da cui risulti un miglioramento della Qualità del prodotto.

5. Conclusioni

La sperimentazione effettuata negli ultimi mesi sui prodotti sopra accennati, che ha portato alle misure dette, ha messo in luce un aspetto della metodologia scelta e applicata che non annulla la validità delle misure e dei concetti esposti, ma ne suggerisce un ampliamento, che rende tale metodologia di interesse e utilità assai più generale.

Il punto nuovo è il seguente.

La difettosità reale di un prodotto in field sarebbe quella misurata secondo l'attuale metodologia se in un dato momento tutti i difetti sino ad allora riscontrati fossero stati conteggiati ed opportunamente pesati, contribuendo cioè alla formula che definisce la Difettosità Totale in Field.

In realtà, in un dato momento esistono di fatto segnalazioni di malfunzionamenti (o failures) non ancora analizzate, che possono poi risultare o causate da effettivi difetti del prodotto, o da altre cause.

Più avanti riprenderemo queste altre cause.

Si può però dire che, mentre questi effettivi difetti non ancora ritenuti tali non entrano nel calcolo della Difettosità Totale a quel momento, e quindi quella misurata risulterà inferiore a quella "reale", l'insieme dei malfunzionamenti verificatisi (indipendentemente se dovuti a difetti del prodotto o ad altre cause) è ciò che dà all'utente la "sensazione" di qualità del prodotto, che potremmo chiamare *"qualità apparente"*. Questa considerazione ha portato ad ampliare il metodo di misura nel senso di conteggiare, nelle formule precedenti, non i *difetti,* ma i *malfunzionamenti.*

Il risultato di tali misure fornisce pertanto, anzichè un *indice di difettosità, una densità di failure* di valore sicuramente maggiore del precedente.

Ciò rappresenta perfettamente il fatto che la qualità recepita dall'utente, che abbiamo chiamato *qualità apparente,* sia inferiore a quella *reale,* mentre da quanto detto prima risulta anche che la *qualità* misurata in base all'*indice di difettosità* sia maggiore di quella *reale*, cioè:

$Q_{APP} < Q_{REALE} < Q_{MIS}$

Per far si che la Qualità misurata fosse uguale a quella reale, bisognerebbe che al momento della misura non esistessero segnalazioni di malfunzionamenti ancora da analizzare (il che dipende dalla dimensione del prodotto, del parco, dall'efficienza del Servizio di Assistenza Clienti e dall'efficienza del gruppo di Manutenzione).

Per far si che la Qualità apparente fosse uguale a quella misurata, o almeno a quella reale, bisognerebbe che non esistessero più malfunzionamenti non dovuti a difetti del prodotto, ossia non esistessero le altre cause cui si era prima accennato.

Quali sono queste altre cause?

Si possono brevemente sintetizzare come segue:

- errore d'uso del sistema da parte dell'utente
- difetto già altrove rilevato, risolto e non prevenuto applicando tempestivamente la correzione
- malfunzionamento di cui non viene fornita documentazione sufficiente ad analizzarlo
- errore apparentemente software, originato da guasto hardware.

La rilevanza di queste cause sul totale dei malfunzionamenti analizzati non è banale, ma varia tra il 40% e il 60%, il che rende significativamente diversi i valori di *Indice di difettosità e densità di failure* e quindi, di conseguenza, i due valori di qualità.

Far si che non esistano le cause sopra citate dipende solo in parte dal prodotto, ma prevalentemente da quanto gli sta intorno, e cioè, formazione, assistenza tecnica, distribuzione prodotti, disponibilità manuali utente, ecc. ecc.

L'archeologia classica e l'elaborazione elettronica dei dati

ROBERTO ROSATI

Istituto di Archeologia
Università di Bologna

1. Introduzione

L'archeologia è una disciplina di confine[1]. Il suo statuto epistemico, già generico in passato (archeologia = disciplina che studia le testimonianze materiali dell'antichità) ha visto dilatare il suo oggetto di studio nello spazio e nel tempo (archeologia di civiltà extra-europee, archeologia medievale, archeologia industriale), al punto che ne appare ormai necessaria una ridefinizione più aggiornata.

Questa indeterminatezza epistemica della disciplina, se da un lato ha favorito la proliferazione dei suoi aspetti più deboli e caduchi legati alle divagazioni antiquarie o erudite che ne hanno in ogni tempo costellato lo sviluppo, dall'altro l'ha resa singolarmente permeabile e recettiva al confronto interdisciplinare e alla sperimentazione di nuove metodologie. L'approccio con una realtà archeologica sempre meno elitaria (il "pezzo" pregiato, esclusivo oggetto di interesse dell'archeologia ottocentesca) ed anzi, al contrario, sempre più estesa quantitativamente; la conseguente dapprima timida, poi sempre più marcata introduzione di metodi statistici e quantitativi; l'adozione, infine, come ulteriore conseguenza, di elaboratori elettronici come indispensabile strumento di analisi di vaste quantità di dati: sono questi gli elementi, propri della archeologia contemporanea, che hanno assestato un duro colpo a tutta una tradizione di studi che, pur nelle differenze inevitabili fra i vari periodi, poggiava nel complesso sulla sistemazione teorica individuale di un quadro archeologico fortemente semplificato e selezionato. Gli apporti interdisciplinari o, comunque, i contatti con altre discipline si sono ulteriormente infittiti, le sperimentazioni di nuove metodologie affollano le più recenti relazioni congressuali: un nuovo paradigma scientifico - per esprimersi secondo la terminologia di Kuhn[2] - sembra coinvolgere anche l'ambito archeologico.

A tale processo di rinnovamento, tuttavia, l'archeologia classica ha contribuito in modo molto esiguo. Mentre l'archeologia preistorica o le archeologie extra-europee, relative a società tipologicamente più semplici di quelle classiche e più libere dalle finalità ideologiche ed estetiche proprie dell'archeologia classica (la normatività della cultura greca e romana e il suo valore nella formazione culturale della borghesia europea), hanno accolto sollecitamente la sperimentazione di metodologie quantitative atte a strutturare e a modellare la varietà dei dati fornita dallo scavo, gli studi di archeologia classica improntati a tale indirizzo di ricerca occupano ancora un posto del tutto marginale nell'ambito della loro disciplina. E non è a caso che dal settore delle archeologie non classiche provengano taluni dei teorici più raffinati della "New Archaeology" o di altre metodologie affini, da Clarke, a Binford, a Renfrew.

Il problema riguarda quella che, nello schema grafico dell'attività archeologica (v. fig. 1) si definisce come fase di elaborazione del modello (iconico, analogico, simbolico)[3] e che, collocandosi dopo le fasi dell'archeologia di campo e di laboratorio, prelude, attraverso ripetuti processi di *feed-back* che rimandano al materiale, alla formulazione di una Teoria finale di sintesi. Mentre infatti, per quanto concerne le fasi relative all'archeologia di campo e di laboratorio, non sussistono differenze strutturali fra l'archeologia classica e le altre archeologie nella raccolta e catalogazione di dati attraverso l'uso dell'elaboratore elettronico (pur nell'ambito delle problematiche e dei criteri tassonomici propri di ogni disciplina), è nella scelta e nella disposizione dei dati in vista della loro elaborazione teorica che l'archeologia classica palesa la sua diversità precludendosi l'accesso a modelli di elevato livello conoscitivo.

È in questa fase che i pur rigorosi procedimenti scientifici della disci-

1) Per l'archeologia come disciplina "frontalière" si v. l'introduzione di A. SCHNAPP al volume di AA.VV., *L'archéologie aujourd'hui,* Paris, 1980, pag. 16. Per uno sguardo d'insieme sulle varie problematiche dell'archeologia attuale appare utile una rassegna dei principali contributi del volume.

2) Il riferimento è, ovviamente, a T.S. KUHN, *The Structure of Scientific Revolutions,* Chicago 1962, tr.it. Torino 1969.

3) Secondo D. CLARKE, *Analytical Archaeology,* London 1968, pag. 446, l'archeologia ha al massimo utilizzato semplici modelli iconici o modelli analogici di basso livello, legati a ipotesi di ordine storico o antropologico. Modelli analogici più sofisticati e modelli simbolici sono in forte ritardo nella disciplina.

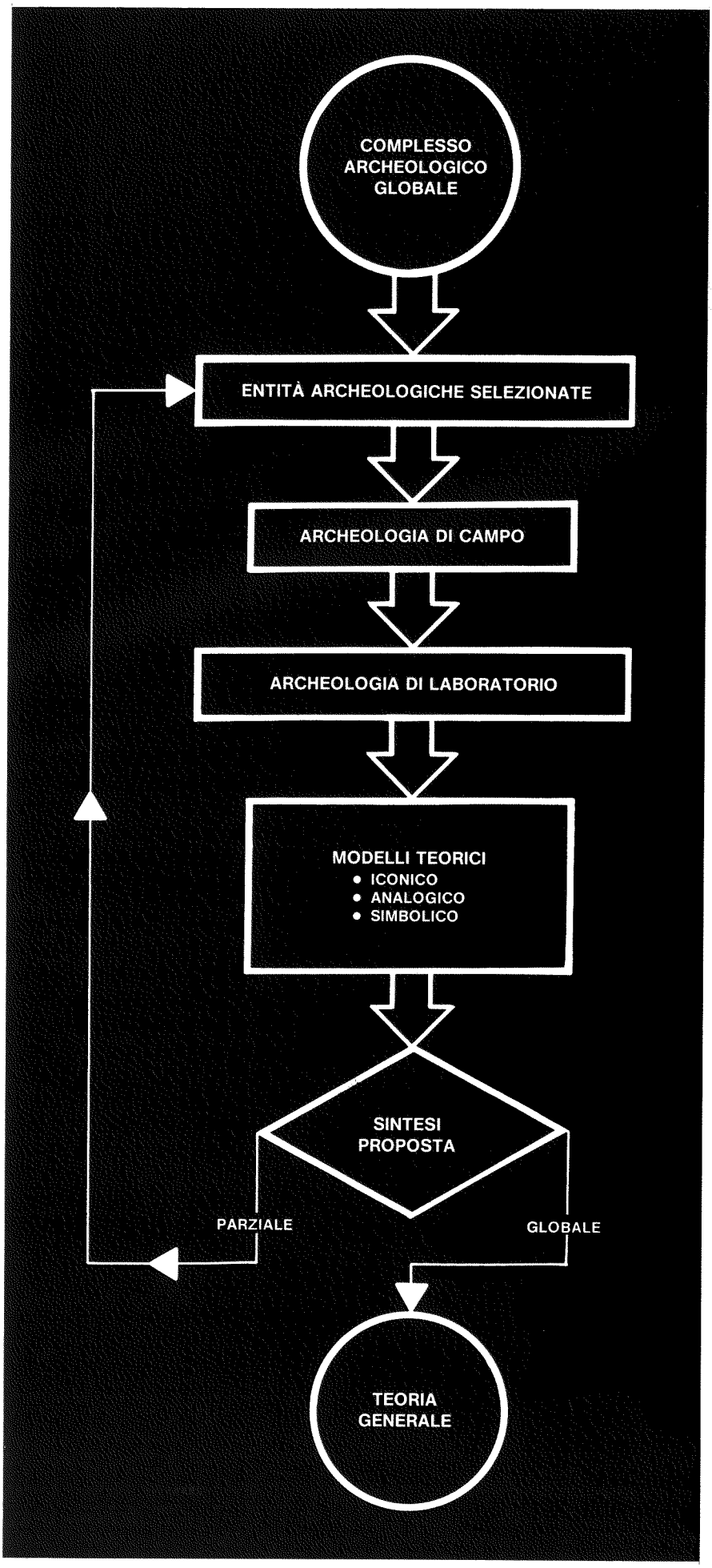


Fig. 1 - Un modello semplificato dell'attività archeologica. Notare il processo di feed-back fra le sintesi provvisorie e il materiale.

plina subiscono un arresto: a questo punto la procedura archeologica prevalente nella disciplina o aggira completamente il problema dei modelli teorici, limitandosi a una esposizione filologica degli esiti dello scavo; o produce delle proposte parziali (e talora anche delle teorie di carattere generale) basate su sintesi di carattere per lo più *narrativo* (legato in genere alle valenze estetiche del materiale e al loro significato per la cultura occidentale). È dunque, in prevalenza, alla sfera relativa alla elaborazione dei modelli e alle possibilità di strutturare anche i dati dell'archeologia classica in rapporto ad esigenze conoscitive di più elevato contenuto teorico, che saranno rivolte le pagine che seguono (le fasi relative all'archeologia di campo e a quella di laboratorio sono state già in precedenza trattate su questa rivista)[4].

2. La "Nuova Archeologia"

Benché oggi si tenda genericamente ad attribuire ad ogni impostazione metodologica basata su criteri quantitativi e/o sull'uso degli elaboratori elettronici l'etichetta di "New Archaeology", in realtà le premesse metodologiche e le modalità operative degli studiosi che aderiscono a tale indirizzo sono alquanto specifiche, e tali da suscitare talora riserve o dissensi da parte di altri studiosi che pure si richiamano a metodologie quantitative[5]. Non ogni elaborazione quantitativa del dato archeologico rientra necessariamente nell'ambito della "New Archaeology", che del resto si è costituita sulla scia di sperimentazioni già da tempo avviate nel settore del trattamento statistico dei dati archeologici. Le prime ricerche in tale ambito vengono di solito ricondotte agli anni Trenta, alle prime sperimentazioni statistiche del Ford[6] legate alla necessità di classificazione di vaste quantità di materiale archeologico in seguito alle attività di scavo condotte negli Stati Uniti al tempo del New Deal rooseveltiano. Il lungo periodo che si estende da queste ricerche pionieristiche alla seconda metà degli anni Sessanta, allorché venne definendosi la "New Archaeology" propriamente detta (e che distinguiamo con le maiuscole), benché non possa fregiarsi di una qualche specifica etichetta, risulta tuttavia caratterizzato da diverse sperimentazioni la cui portata dovrà un giorno essere rivalutata in rapporto alle più coerenti teorizzazioni del periodo successivo. È in questa fase che, ad opera di taluni studiosi che fanno per lo più riferimento alla rivista *American Antiquity,* vengono approfonditi i due versanti di maggior rilevanza teorica ai fini dell'applicazione di metodi quantitativi al materiale archeologico: quello tassonomico, volto alla definizione e classificazione delle entità individuate, e quello più propriamente statistico. Al primo versante saranno da ascrivere alcuni lavori rivolti al concetto di "tipo" o di "artefact" o di "artefact-type" da parte di studiosi come Krieger[7], Ford[8], Rouse[9], mentre al secondo è rivolto lo studio di Spaulding[10] sulle tecniche statistiche nella definizione degli "artefact-types", nonché il contributo di Heizer e Sherburne[11] sulla applicazione di metodi quantitativi in archeologia, che costituisce una prima provvisoria *mise-au-point* degli sforzi e dei progressi delle teorizzazioni del settore, dall'impresa pionieristica di Ford fino agli inizi degli anni Sessanta. È da rilevare che proprio in quegli anni appare per la prima volta la consapevolezza che l'archeologia ha imboccato degli indirizzi metodologici assolutamente specifici e originali, con l'esplicito riferimento alla esistenza di una "new archaeology" (il termine è ancora usato genericamente e non in riferimento ad una scuola specifica) che appare in un articolo di J. R. Caldwell sulla rivista "Science", apparso nel 1959[12].

L'attività di Gardin in Francia, mentre completa da un lato con alcuni contributi teorici apparsi nella prima metà degli anni Sessanta taluni temi sviluppati nel corso del dibattito precedente[13], avvia anche concretamente il problema della realizzazione di talune strutture (ad. es. la creazione di una codifica delle forme ceramiche) necessarie al trattamento informatico dei dati archeologici[14]. Non esiste, ovviamente, un atto di nascita ufficiale della New Archaeology. Non dovrebbero però sussistere eccessive difficoltà nell'attribuire al 1968 il ruolo di anno cruciale nella formulazione delle teorie della scuola. È in quest'anno, infatti, che appaiono due dei maggiori contributi teorici di esponenti della scuola: il famoso *Analytical Archaeology* di David Clarke[15] e *New Perspectives in Archaeology*, curato dai Binford[16]. È da questi due contributi che trarremo gli elementi essenziali per una rapida caratterizzazione dei tratti distintivi di questa scuola.

4) Cfr. C. FALCETTI, *Elaboratori e archeologia,* sul n. 23, XI, (1984) di questa rivista. pag. 37 e ss.

5) Cfr. ad es. J.E. DORAN - F.R. HUDSON, *Mathematics and Computers in Archaeology,* Edinburgh (1975), pag. 5. Più favorevole invece il giudizio di S. CLEZIOU, J.P. DEMOULE, A. SCHNAPP, *Renouveau des méthodes et théorie de l'archéologie,* in *Annales* ESC, (1973), pp. 35 ss.

6) J.A. FORD, *Ceramic decoration sequence at an old indian village,* New Orleans, 1935.

7) A. KRIEGER, *The typological Concept,* in *American Antiquity,* 5, (1944), pp. 271 ss.

8) J.A. FORD, *The type concept revisited,* in *American Anthropologist,* 56 (1954), pp. 42 ss.

9) I. ROUSE, *The Classification of artefacts in Archaeology,* in *American Antiquity,* 25, (1960), pag. 313 ss.

10) A.C. SPAULDING, *Statistical techniques for the discovery of artefact types,* in *American Antiquity,* 18 (1953), pag. 305 ss. Cfr. anche D. de SONNEVILLE-BORDES, J. PERROT, *Essai d'adaptation des méthodes statistiques au Paléolitique supérieur. Premiers resultats,* in *Bull. Soc. Prehist. Franç.,* 50 (1953), pp. 323 ss.

11) R.F. HEIZER - F.C. SHERBURNE (edd.), *The Application of quantitative methods in Archaeology,* in *Viking Fund Publications in Archaeology,* 28 (1960).

12) J.R. CALDWELL, The *New American Archaeology,* in *Science,* 29, (1959), pp. 303 ss.

13) J.C. GARDIN, *Problèmes d'analyse descriptive en archéologie,* in P. COURBIN (ed.), *Études archéologiques,* Paris (1963), pp. 132 ss.; ID., *Analyse documentaire et analyse structurale en archéologie,* L'Arc, Aix-en-Provence, 26 (1965), pp. 64 ss.; ID., *On a possible interpretation of componential analysis in archaeology,* in E.A. HAMMEL (ed.), *Formal semantic analysis, American Anthropologist,* special pubblication, 67 (1965), pp. 9 ss; ID., *Methods for the descriptive analysis of archaeological materials,* in *American Antiquity,* 32, 1, (1967), pp. 13 ss.

14) J.C. GARDIN, *Code pour l'analyse des formes de poteries,* Paris CNRS 1976.

15) D. CLARKE, op.cit. (a nota 3).

16) S.R. BINFORD - L.R. BINFORD, *New Perspectives in Archaeology,* Chicago 1968.

Come si è già osservato, la valutazione culturale della New Archaeology deve considerare i suoi propositi specifici e non la semplice propensione all'uso dell'elaboratore elettronico o dell'elemento quantitativo. Sotto questo profilo, benché tale indirizzo non costituisca certamente la totalità di quanto di nuovo si sta attualmente elaborando nel campo dell'archeologia, e i suoi postulati teorici possano anche non essere agevolmente accettati senza riserve, occorre però riconoscere come essa abbia per la prima volta raccolto il frutto della discussione che avveniva in ordine sparso negli anni Cinquanta e nei primi anni Sessanta, ed abbia disposto entro una struttura che può essere discutibile ma che appare teoricamente robusta tutto un complesso di osservazioni e di sperimentazioni che rischiava di frammentarsi.

Il primo postulato fondamentale della New Archaeology fa riferimento al dibattito in corso sui fondamenti epistemologici delle scienze (teorie di Kuhn, Hempel etc.) e propone come preliminare il carattere *deduttivo* della ricerca scientifica[17]. Ironizzando sui principi induttivi propri delle ricerche archeologiche tradizionali (che assimilano il lavoro dell'archeologo a una raccolta indiscriminata di dati, così come essi si offrono alla sua attenzione, e da cui poi viene aggregandosi la sua teoria) si insiste sulla priorità che assume la teoria o il progetto, che deve essere già formulato dal ricercatore ed essere operante quando questi passa a selezionare i dati (che sarebbero altrimenti di per sé infiniti) della realtà archeologica. Il seguito del procedimento scientifico consiste in un continuo processo di *feed-back* fra la teoria e la verifica empirica nella realtà archeologica concreta.

Merito di David Clarke è stato di aver offerto, col suo *Analytical Archaeology* (il famoso Clarke bleu, dal colore della copertina) un insieme coerente di teorizzazioni entro il quale inserire il vasto complesso delle tecniche matematiche e statistiche. Esigenza primaria di ogni disciplina scientifica è - secondo Clarke - la definizione preliminare delle entità su cui opera e dei processi e delle relazioni che collegano tali entità, così come avviene ad es. in geometria per il punto, la retta, il piano[18]. Lo studioso ironizza al riguardo sul carattere di "undisciplined discipline" che assume attualmente l'archeologia, sul suo aspetto narrativo, sulla frequenza dei giudizi di valore. Alla base della struttura gerarchica delle entità archeologiche che l'autore pone l'*attribute*[20] la più piccola unità informativa non ulteriormente scomponibile, costituita da qualunque carattere segnico impresso dall'uomo sul materiale. Un complesso di *attributes*[21] (preventivamente selezionati dal ricercatore fra il vasto numero quelli a lui ritenuti più indicativi) costituisce l'*artefact* (che qui per semplicità fondiamo con l'*artefact type*), denominazione sotto la quale possono rientrare numerosi tipi e oggetti noti all'archeologia tradizionale (anfora, cratere, kouros etc.). Un complesso di *artefacts* la cui associazione non sia meramente additiva ma dotata di significato, costituisce l'*assemblage* (un semplice esempio di *assemblage* è riportato in fig. 2). Un complesso di *assemblages* genera infine la *culture*, che costituisce ormai un' entità archeologica di scala regionale. Entità più complesse, come il *culture group* o il *technocomplex*, ma dotate di minor potenziale informativo, vengono usate più raramente.

Alla descrizione analitica e ai processi dinamici di tali entità, che nel loro complesso veicolano l'informazione nell'ambito del sotto-sistema della cultura materiale, è dedicato ampio spazio nel volume di Clarke. Il complesso della cultura materiale, definito dalle entità descritte, si integra con altri sotto-sistemi (mitico-religioso, psicologico etc.) in modo da disciplinare i flussi informativi che regolano la gestione del macrosistema socio-culturale, che li comprende e ne organizza le relazioni con l'ambiente esterno (clima, flora, fauna, altri sistemi socio-culturali) (v. fig. 3)).

Esplicita è nel Clarke l'affermazione di una possibile capacità predittiva dello sviluppo del sistema nel tempo, legata alla possibilità di configurare una matrice di risposte probabilistiche fornite dal sistema al variare degli input. (v. fig. 4).

La fase attuale di ricerca sembra caratterizzata da una maggior cautela circa gli obiettivi finali piu ambiziosi di una gestione quantitativa dei dati archeologici, e da una maggiore attenzione, per contro, all'approfondimento dei singoli aspetti tecnici delle diverse sperimentazioni quantitative. Così, talora nell'ambito della New Archaeology, talora al di fuori di essa, continuano le diverse ricerche nei più svariati settori del trattamento quantitativo dei dati. Dall'ambito strettamente statistico, ove vengono approfondite le diverse applicazioni alle realtà archeologiche delle varie tecniche (test del *chi* quadro, deviazione standard, rette di regressione, analisi multivariata)[22]; a quello della *spatial archaeology*, un ramo che sta approfondendo autonomamente le proprie ricerche occupandosi dell'identificazione di aggregati archeologici *non-random* nello spazio attraverso l'uso di appropriate metodologie matematiche ("nearest neighbour analysis", e analisi multivariata)[23]; a quello delle simulazioni, un settore di cui un recente volumetto propone talune applicazioni riguardanti nel complesso società dalla struttura piuttosto semplice, e che appare di difficile trasposizione nell'ambito delle società classiche[24].

17) Per una discussione aggiornata e per un'esplicita posizione a favore del metodo deduttivo cfr. ora J.N. HILL, *The methodological debate in contemporary archaeology: a model,* in D. CLARKE (ed.), *Models in archaeology,* London 1972, pag. 61 ss.

18) D. CLARKE, *Analytical Arch.,* cit. pag. 14 ss.

19) ID., *ibid.,* pag. XIII.

20) ID., *ibid.,* pag. 131 ss.

21) Queste successive aggregazioni di entità archeologiche vengono realizzate dal Clarke su base politetica e non monotetica, vale a dire in base al possesso di una certa quantità (non rigida) di proprietà. L'adozione di un rigido criterio monotetico (presenza necessaria e sufficiente di certe determinate proprietà per l'appartenenza ad una categoria) limiterebbe eccessivamente la flessibilità naturale implicita in tali associazioni. Tali criteri sono stati ripresi dal Clarke da SOKAL & SNEATH, *Principles of numerical taxonomy,* San Francisco-London, 1963.

22) Per un quadro sintetico e chiaro delle ricerche in corso in tale ambito cfr. A. BIETTI, *Tecniche matematiche nell'analisi dei dati archeologici,* Roma (Accademia Nazionale dei Lincei), 1982.

23) Cfr. I. HODDER - C. ORTON, *Spatial analysis in archaeology,* Cambridge 1976. Si v. anche R. WHALLON Jr., *Spatial analysis of palaeolithic occupation areas,* in C. RENFREW (ed.), *The Explanation of Culture Change,* Liverpool 1973, pag. 115 ss.

Fig. 2 - Un semplice *assemblage* di età geometrica. Un grande vaso funerario (per lo più un cratere o un'anfora) emerge dal terreno come segnacolo. Sotto, il cinerario. Talora presenti, come corredo, piccoli vasi geometrici. (da Perrot-Chipiez).

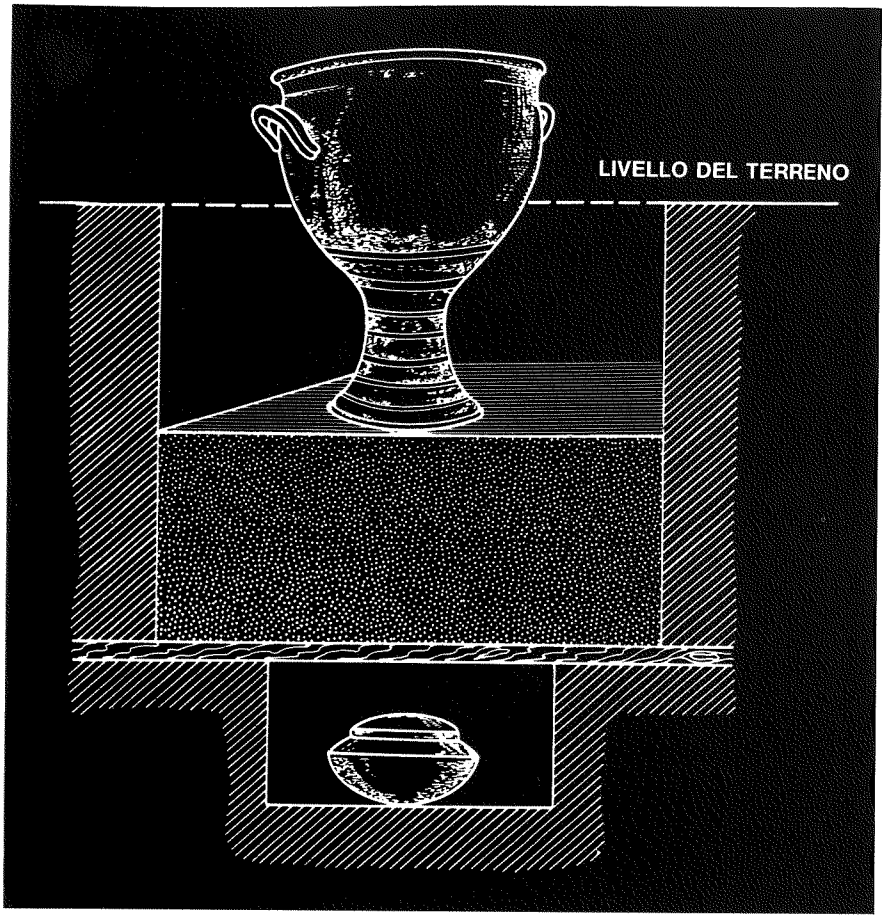
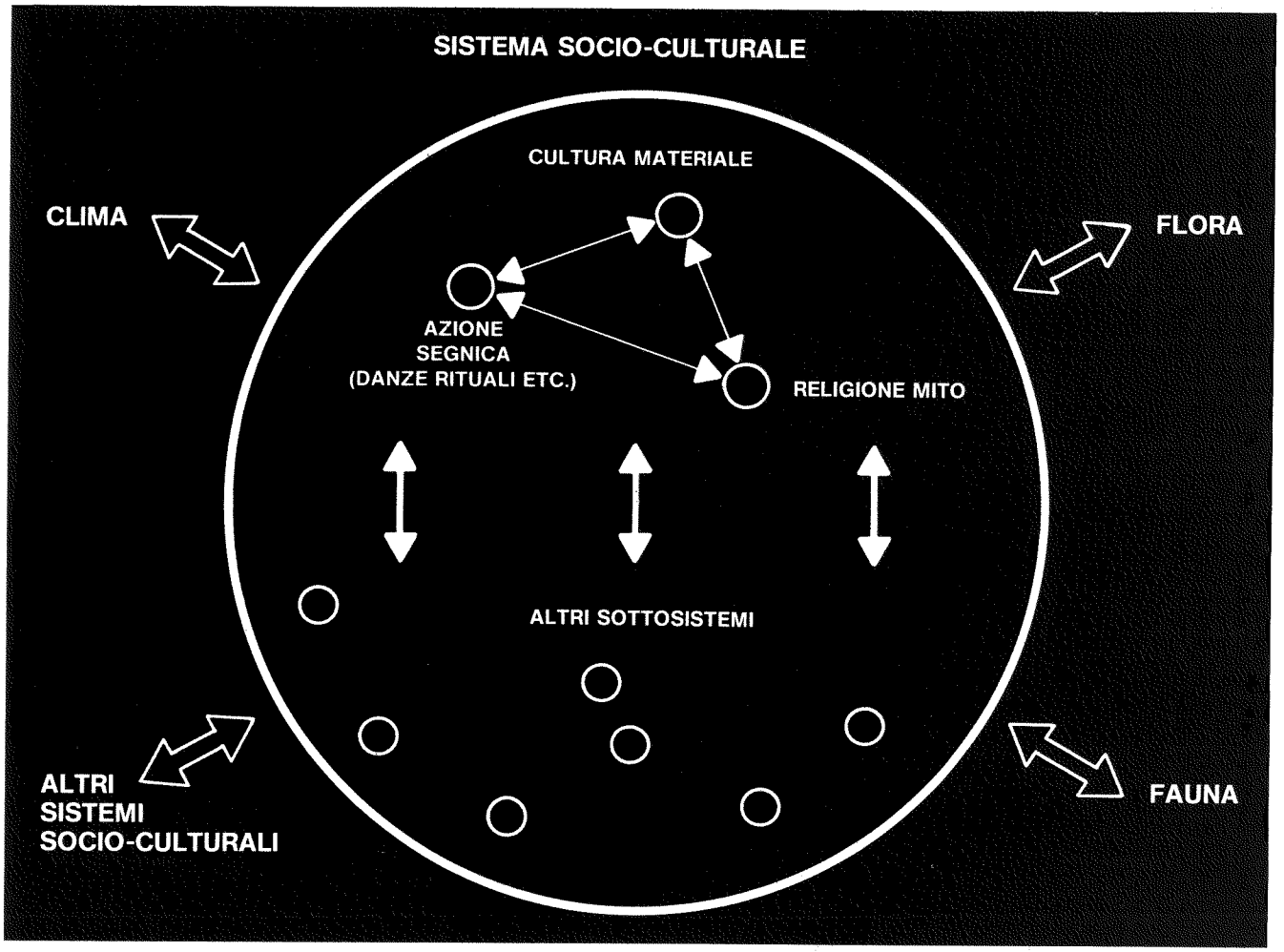


Fig. 3 - Il sistema socio-culturale secondo D. Clarke (da Analytical Archaeology, London 1968, rielaborato).



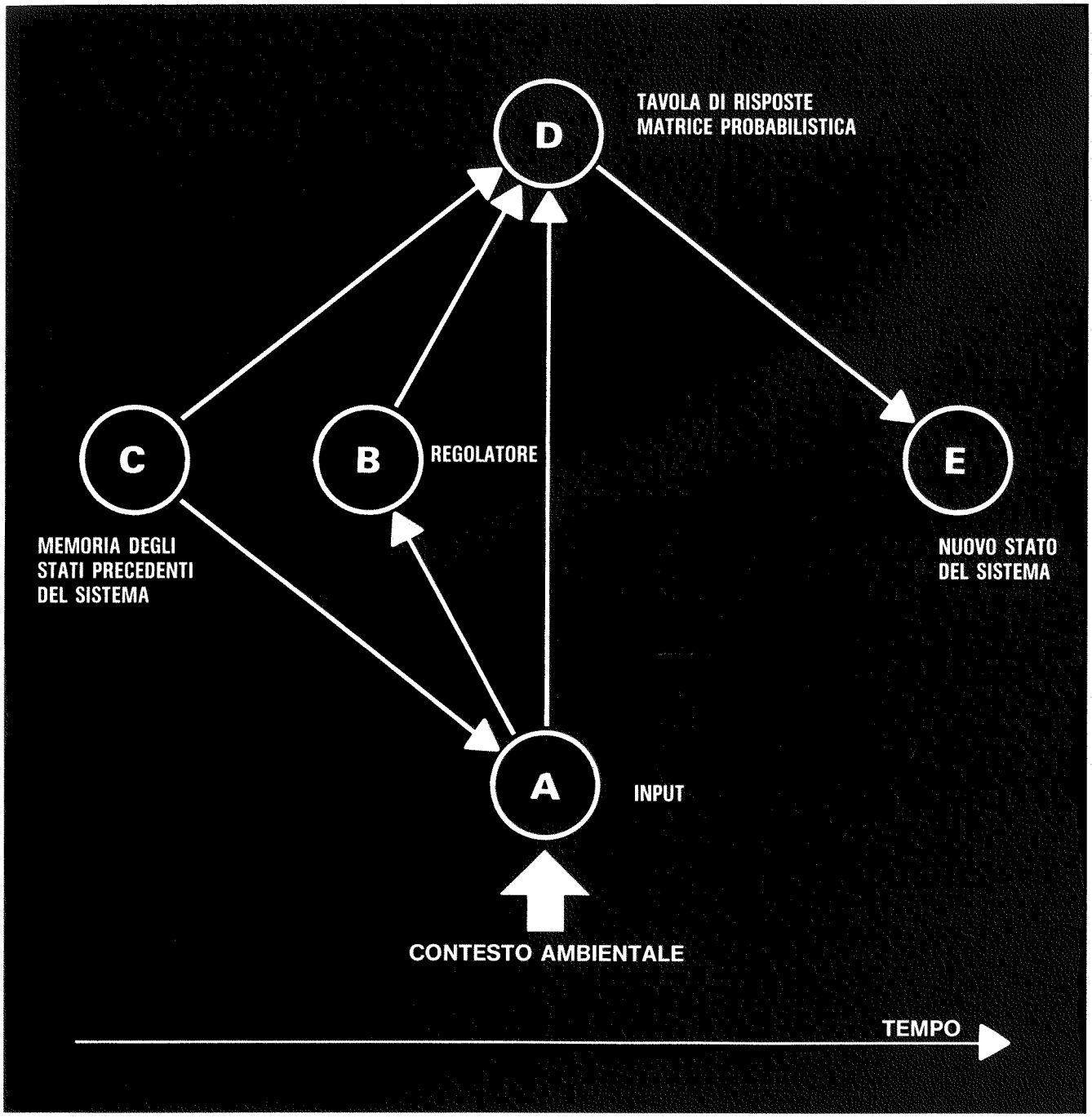


Fig. 4 - L'evoluzione di un sistema socio-culturale secondo D. Clarke (da Analytical Archaeology, London 1968).

3. Elaborazione di modelli

Una tradizione culturale quale quella della *Altertumswissenschaft,* la scienza filologica dell'antichità classica, disciplina guida del campo umanistico nel secolo scorso e in parte dell'attuale, saldamente radicata nel patrimonio culturale europeo, non appare certo la più idonea a promuovere innovazioni metodologiche. L'obiettivo di tradurre i suoi dati in forme assimilabili dalle griglie interpretative di modelli di elevato valore teorico sembra in effetti un'impresa difficile, di fronte a procedimenti e obiettivi culturali ormai consolidati. Mentre è in corso da tempo, nelle riviste del settore, il dibattito sui caratteri che deve assumere il rinnovamento, si può solo accennare, in questa sede, alla possibile evoluzione di questi studi verso una dimensione meno qualitativa[25], in direzione delle potenzialità più vaste, di ordine conoscitivo, dischiuse da tematiche più moderne come quelle relative all'informazione e alla storia economico-socia-

24) I. HODDER (ed.), *Simulation Studies in Archaeology,* Cambridge 1978.

25) In tal senso, indicativa di un'aspirazione dell'archeologia classica verso un ampliamento alla cultura materiale appare il contributo di G. CARANDINI, *Archeologia e Cultura Materiale,* Bari 1975 (ried. ampliata Bari 1979). Il mancato riferimento a metodi quantitativi per fronteggiare il più vasto materiale, circoscrive però sensibilmente la portata concreta della proposta.

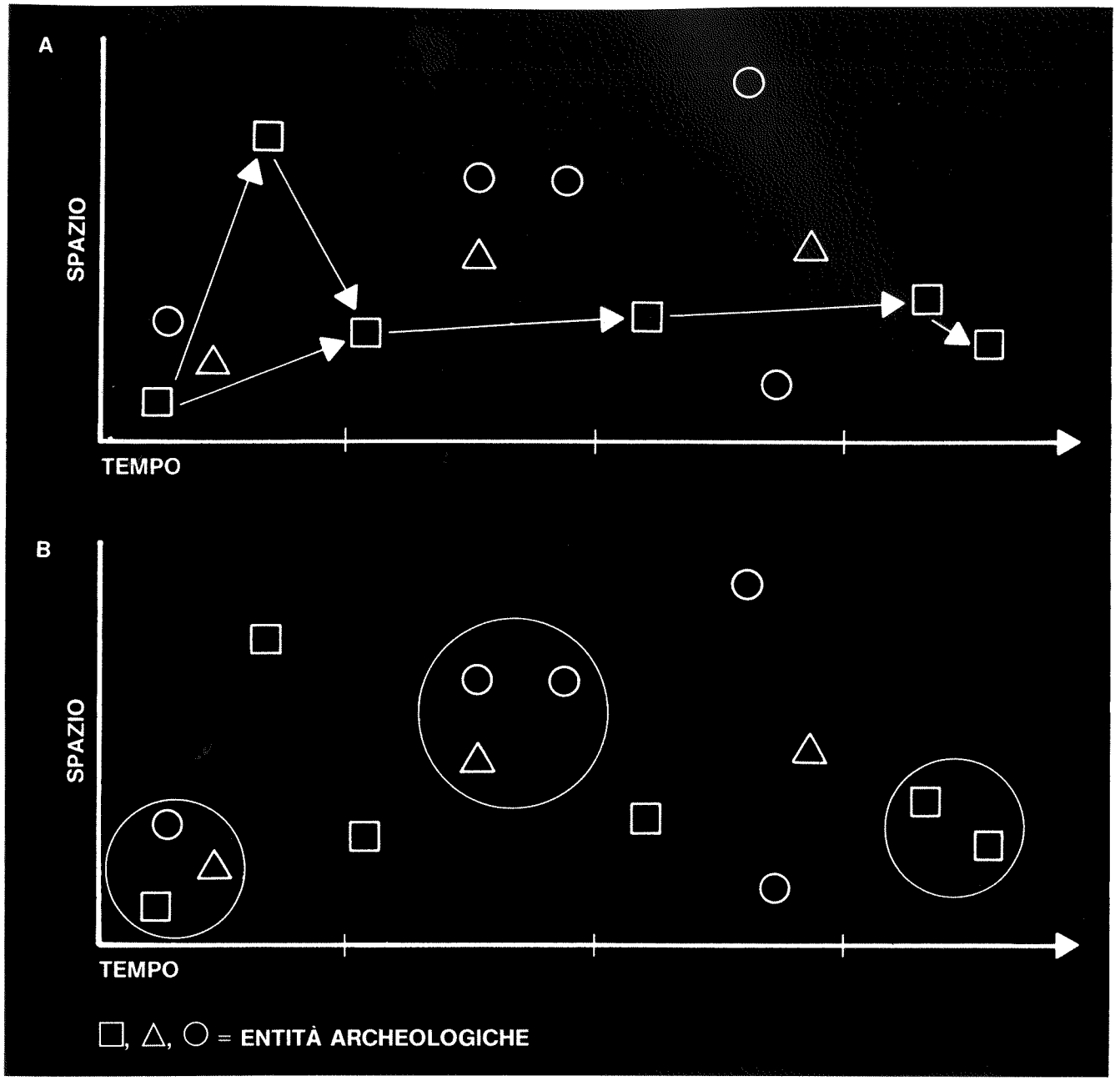


Fig. 5 - La procedura scientifica tradizionale dell'archeologia classica (A) e la moderna ricerca di *assemblages* (B).

le. L'obiettivo di una formalizzazione avanzata del trattamento dei dati resta ancora lontano, ma il passo è indispensabile se si vuole procedere verso modelli di elaborazione teorica dei dati che non siano solo sintesi individuali di carattere intuitivo.

Pur nella presenza di talune specifiche difficoltà, un programma di ricerca relativo al tema dell'informazione veicolata dalla cultura materiale come sotto-sistema di un sistema socio-culturale può sostenere un vasto ambito di studi e di sperimentazioni particolari anche nell'ambito dell'archeologia classica. La Polis greca, così come, del resto, la città medievale o lo stato moderno, costituisce un sistema socio-culturale, una particolare forma di sistema aperto suddivisa in un complesso di sotto-sistemi intercorrelati da flussi di informazione e alimentato dal prelievo di energia dall'esterno. A una molteplicità di compiti attendono i sotto-sistemi preposti all'informazione[26] che, in sintesi, regolano l'attività delle diverse componenti del sistema e assicurano la continuità del suo funzionamento, almeno in riferimento a margini di variazione tollerabili dell'input. Fra tali sottosistemi preposti all'informazione il Clarke distingueva come più importanti, nelle società di tipo arcaico, quelli rela-

26) Fra tali compiti: "Scambiare dati e messaggi fra due sistemi, sviluppare e formare la conoscenza, inviare comandi di regolazione e controllo, organizzare la struttura interna di un sistema, esercitare il potere". Così G.B. e M.L. SCURICINI, *Cibernetica e noetica,* Firenze 1985, pag. 111.

tivi alla sfera dell'azione (segnica), della tradizione, della cultura materiale[27].

L'approccio sistemico nel settore dei processi culturali, se consente un quadro sicuramente più moderno rispetto alla prospettiva storicistica tradizionale, soffre tuttavia di alcune limitazioni, almeno se applicato all'ambito archeologico[28]. In effetti, talune ottimistiche impostazioni della fine degli anni Sessanta e dei primi anni Settanta sembrano al momento ridimensionate: così ad es. la pretesa del Clarke di delineare la traiettoria di un sistema socio-culturale nel tempo. Gli studi più recenti escludono, per talune classi di sistemi, la possibilità di predeterminare una direzione[29]; la teoria delle catastrofi, che si occupa degli improvvisi mutamenti di stato nei sistemi, non ha al momento applicazioni concrete; il concetto stesso di informazione, infine, nella misura in cui viene approfondito, si rivela complesso e dotato di componenti non facilmente riducibili all'analisi quantitativa. In effetti l'informazione studiata da Shannon e Weaver[30], quantificabile nella misura in cui riduce l'incertezza del ricevente, costituisce solo *uno* degli aspetti di questa complessa entità se esaminata nei contesti sociali. Meno facilmente quantificabili appaiono gli aspetti qualitativi del fenomeno, legati al problema del *significato*, cioè del valore che un messaggio assume per un sistema, nonché gli aspetti *organizzativi*, che pure costituiscono un altro dei compiti dell'informazione nella sua concreta funzione sociale[31].

Queste restrizioni devono essere rilevate al fine di delimitare l'effettiva attuabilità di un progetto di ricerca sull'informazione come elemento di regolazione nell'ambito dei sistemi socio-culturali. Il progetto, in effetti, va pensato non come già preordinato in ogni sua parte, ma come scoperta e progressiva estensione di ambiti più limitati di ricerca in cui metodologie matematiche e statistiche possono essere utilmente adottate in modo non occasionale. Il confronto e la progressiva integrazione di tali ricerche contribuirà a una graduale estensione degli ambiti formalizzati e ad una successiva armonizzazione degli obiettivi. Entro questi limiti e sulla base di tali criteri di attuazione del progetto, è possibile prospettare anche per l'archeologia greca e romana un orientamento di ricerca suscettibile di un trattamento teorico dei dati anche da parte dei modelli più sofisticati.

L'archeologia classica tradizionale, date le sue prevalenti finalità estetiche, ha finora nettamente privilegiato un discorso condotto per singole tipologie, studiate genealogicamente, dalle origini alla loro estinzione, nella loro diffusione spaziale (es. l'origine e lo sviluppo del frontone: (v. fig. 5). Il compito che si apre ora è invece: a) ricostruire il modo in cui le entità archeologiche si aggregano nel territorio in complessi dotati di senso b) studiare le funzioni esercitate da tali strutture comunicative nella gestione complessiva dell'informazione. Un simile programma di lavoro, almeno nella sua forma più essenziale, dovrebbe prevedere:

1) L'individuazione (da parte dell'archeologo) delle unità di base (gli *artefact-types*) che costituiscono le entità costitutive delle strutture comunicative cercate (gli *assemblages*).

2) La ricostruzione della rete che correla gli *assemblages* di una regione o territorio, il che equivale a ricostruire la sua *culture*, cioè il trattamento globale dell'informazione veicolata dalla cultura materiale.

   (Particolarmente utile appare in queste 2 fasi l'uso dell'elaboratore elettronico per collocare le singole unità di base in un reticolo spaziale e riconoscere aggregati significativi di oggetti ed eventualmente compararli).

3) Lo studio, infine, delle variazioni di entropia delle strutture comunicative identificate, al fine di seguire, nel corso del tempo, la loro maggiore o minore rilevanza nell'ambito delle ristrutturazioni periodiche della rete informativa che il sistema effettua in relazione al variare degli input energetici.

Esemplifichiamo concretamente.

Supponiamo di dover condurre, sulla base dei criteri appena enunciati, un esame della cultura materiale, intesa come sistema di informazioni, della Polis attica, cioè di Atene e del suo territorio.

Un primo esame anche superficiale della regione rivelerà immediatamente taluni *assemblages* particolarmente appariscenti, come l'acropoli o l'agorà. In entrambi i casi siamo in presenza di strutture comunicative particolarmente enfatizzate (per qualità e valore economico dei materiali) dal sistema di cui fanno parte, e ciò non è senza rilevanza per l'individuazione della gestione sociale dei flussi energetici che irrorano il sistema: in effetti, il modo in cui questi flussi vengono organizzati socialmente (o politicamente) lascia una traccia significativa nell'organizzazione delle strutture comunicative. Così in età classica, quando è attraverso la struttura politico-ideologica della Polis (e non più attraverso le singole casate aristocratiche) che il sistema si procura e redistribuisce all'interno una parte cospicua delle sue risorse energetiche, questa nuova disposizione determina tutta una riorganizzazione, rispetto all'età arcaica, delle strutture comunicative del sistema che fa deperire i vecchi *assemblages* legati ai santuari locali e ai culti funerari gentilizi a tutto vantaggio di quelli legati alla Polis (acropoli, agorà).

Accanto a questi macro-assemblages ne esistono altri minori, che possono ancora essere rintracciati a vista: cimiteri di casate aristocratiche, costituiti dalla presenza di grandi statue e stele funerarie; santuari urbani ed extraurbani dotati di materiale votivo. Una analisi condotta su un reticolo piuttosto dettagliato e con l'ausilio di un elaboratore elettronico può consentire di integrare e reperire quelle ulteriori

27) D. CLARKE, *Analytical* etc., *cit.;* pag. 101

28) Cfr. R.J.C. MUNTON, *System analysis: a comment,* in C. RENFREW (ed.), *The Explanation* etc., cit., pag. 687, che colloca i sistemi archeologici nell'ambito di una categoria di sistemi (morfologici) la cui conoscenza è limitata al solo ambito delle relazioni reciproche all'interno del sistema.

29) Per una rapida rassegna delle più recenti posizioni di Haken, Thom, Prigogine cfr. ancora G.B. e M.L. SCURICINI, *op. cit.,* pag. 90.

30) C.E. SHANNON - W. WEAVER *The Mathematical Theory of Communication,* 1949, tr. it., Milano 1975.

31) Per questi diversi aspetti dell'informazione nei sistemi v. ancora G.B. e M.L. SCURICINI, *op. cit.,* pag. 122.

associazioni di *artefacts* meno appariscenti o meno studiate, che sono in realtà tanto più importanti quanto più diffuse capillarmente nel territorio. Appare importante, in questa fase, l'adozione di tecniche di "cluster analysis" o di analisi multivariata, che consentano di definire e apparentare fra loro i diversi complessi di materiali.

Ricostruita la rete delle strutture comunicative relative alla cultura materiale, un elemento importante della ricerca è dato dalla valutazione delle variazioni di entropia attraverso il tempo. In effetti se, come si è detto, la dislocazione di questa rete comunicativa nel territorio può aiutare a ricostruire l'organizzazione sociale dei flussi di energia che alimentano il sistema, una evoluzione delle caratteristiche informative di una fonte, che varii nel tempo la sua entropia, può suggerire considerazioni interessanti sul suo costituirsi come struttura comunicativa e sul suo degradare. Così è interessante il caso dell'Acropoli, che risorge come struttura comunicativa (non come entità fisica, beninteso!) all'inizio del VI secolo, dopo l'abbandono succeduto al rovinoso crollo della società micenea. L'entropia di tale fonte decresce progressivamente fino a tutta l'età classica, per poi risalire rapidamente in età ellenistica, quando all'alimentazione del sistema diventano del tutto accessorie e superate l'organizzazione della Polis e le strutture anche comunicative ad essa legate.

Se il degrado (la crescita di entropia) di una struttura comunicativa può anche essere colto intuitivamente, resta complesso il problema di una sua esatta quantificazione, problema legato, come già accennato, alla natura composita e ricca di aspetti qualitativi dell'informazione. Al riguardo si può solo avanzare qualche considerazione che contempli la natura eterogenea di strutture comunicative come gli *assemblages*. Qualora, ad esempio (ma il caso è raro) il potenziale di informazione di tali strutture dipenda da una precisa collocazione dei suoi elementi entro una ben precisa sintassi o disposizione spaziale, sarebbe sufficiente valutare gli scarti che si verificano, nel corso del tempo, dalla struttura o disposizione sintattica canonica per constatare l'accrescersi del disordine e quindi l'aumento di entropia. Piu utile e più frequente può essere il ricorso alle metodologie matematiche (tipo "nearest neighbour analysis") usate per la individuazione di strutture di materiali *non-random:* in questo caso si tratterrebbe di seguire, nel corso del tempo, il costituirsi di queste strutture *non-random* e constatarne successivamente la tendenza verso caratteristiche di sempre maggiore casualità di organizzazione interna. Tale procedimento appare però più proficuo per le archeologie preistoriche, ove l'informazione appare meno ricca di elementi qualitativi. La realizzazione di determinati *assemblages,* nel mondo classico, come ad es. l'acropoli di Atene pone non soltanto il problema della coesistenza di determinate tipologie, ma anche del valore intrinseco del materiale (marmo, oro, avorio) o del valore artistico: tutti elementi che accrescono il potenziale informativo. La graduale scomparsa, ad es., di offerte votive di elevato valore artistico o costituite di materiale prezioso o di grandi dimensioni assumerebbe un ben preciso significato nel senso di una variazione verso l'aumento di entropia. In questo caso sembra opportuno tenere in considerazione la presenza di diversi indici, alcuni forniti di una scala di parametri, altri connotati dal semplice tratto presenza/assenza, al fine di stabilire dei criteri sufficientemente comprensivi di valutazione.

4. Gestione dei dati archeologici mediante computer.

Meno complessa sul piano teorico appare la gestione quantitativa dei dati archeologici legata ad un altro aspetto del rinnovamento dell'archeologia classica, orientato in questo caso verso i problemi economico-sociali delle società antiche. Tanto più prezioso giunge in questo caso il contributo dell'archeologia perché in grado di supplire alla lacunosità delle fonti letterarie fornendo preziosi indici quantitativi che assolvono le stesse funzioni svolte in altre civiltà molto più recenti dalla conservazione di documenti scritti[32].

Sia consentito, in questo caso, fare riferimento alla concreta attività di ricerca svolta dall'Istituto di Archeologia di Bologna in collaborazione con il C.I.T.A.M. (Centro di Telematica dell'Università) su elaboratore Honeywell 6/20. Una ricerca avviata nel 1984 e che ha come scopo di costituire una banca dati relativa alla esportazione della ceramica attica[33], rendendone possibile la consultazione a seconda delle principali voci di interesse scientifico (es. provenienze, soggetti, autori epigrafi, forme etc.) anche combinate fra loro. Al momento è stata completata una prima *tranche* relativa a circa 2000 pezzi e riguardante il periodo iniziale di esportazione della ceramica attica a figure nere (630-550 a.C.).

La ricerca è stata avviata in base alla considerazione che solo l'elaboratore elettronico può illuminare certi aspetti del fenomeno commerciale dell'esportazione della ceramica attica: e questo per l'assoluta predominanza del dato quantitativo, che ha sempre scoraggiato, almeno fino all'avvento dell'età degli elaboratori elettronici, ogni possibile sviluppo delle indagini in tal senso (ad eccezione di qualche incompleto e volenteroso tentativo condotto solo per singole chiavi di ricerca).

È impossibile qui rendere conto della vastità delle correlazioni ottenute e dell'importanza dei risultati, sia pur relativi a questa serie di dati: basti solo pensare alla possibilità di ottenere in tempi rapidissimi risposte a domande che hanno a lungo interessato gli studiosi del settore, come ad es. il rapporto fra soggetti e determinate località di esportazione, ovvero fra certe officine attiche e determinati mercati, oppure ancora alla preferenza di determinate località per talune forme.

Nell'impossibilità di rendere conto in dettaglio di tutti i molteplici aspetti della ricerca (una pubblicazione scientifica è in corso fra i suoi curatori), si riportano qui solo alcuni dati scelti fra quelli di interesse più generale. In primo luogo vengono riportati i risultati relativi all'andamento del commercio attico nel

32) Rimando in proposito al mio articolo *Archeologia e Storia dell'Antichità Greca e Romana,* in *Studi Storici, XXI,* 1 (1980) pag. 157 ss.

33) Ad essa hanno collaborato, con lo scrivente, le drr. Maria Pia Guermandi e Lucia Quartili, che qui ringrazio anche per le utili informazioni fornitemi nella stesura di queste righe. Ringrazio inoltre i drr. Carlo Polacchini e Daniele Barbieri del C.I.T.A.M. per i cortesi suggerimenti di ordine informatico.

ESPORTAZIONE CERAMICA ATTICA DAL 630 AL 550 A.C.

	630/600 a.C.	600/575 a.C.	575/550 a.C.
GRECIA	93,84 %	65,66 %	47,27 %
AFRICA	3,08 %	13,05 %	9,80 %
ITALIA	1,54 %	14,06 %	34,24 %
AREA PONTICA	1,54 %	2,01 %	1,82 %
ASIA MINORE	--	2,61 %	1,72 %
GALLIA MERID.	--	1,61 %	4,34 %
HISPANIA	--	1,00 %	0,81 %

Fig. 6 - L'andamento complessivo dell'esportazione della ceramica attica fino alla metà del VI secolo a.C.

ESPORTAZIONE DELLE COPPE E DELLE ANFORE ATTICHE (dal 600 al 550 a.C.)

	COPPE (Kylikes)		ANFORE	
	600/575a.C.	575/550a.C.	600/575a.C.	575/550a.C.
GRECIA	48,48 %	44,85 %	61,22 %	17,24 %
ITALIA	33,33 %	41,78 %	27,56 %	63,55 %
AFRICA	12,12 %	7,52 %	4,08 %	6,90 %
AREA PONTICA	4,55 %	2,23 %	--	0,98 %
ASIA MINORE	1,52 %	1,67 %	4,08 %	1,97 %
GALLIA MERID.	--	1,39 %	3,06 %	8,87 %
HISPANIA	--	0,56 %	--	0,49 %

Fig. 7 - Un confronto fra l'esportazione delle coppe e delle anfore nella prima metà del VI sec. a.C.

bacino del Mediterraneo nei tre periodi principali in cui è stata suddivisa la ricerca (630-600; 600-575; 575-550 a.C.) (v. fig. 6).

Come appare evidente, dopo un primo periodo di semi-isolamento commerciale (630-600) dovuto alla fase di isolamento politico e culturale dell'età presoloniana, in cui la maggior parte della produzione appare destinata al mercato interno, si accentua nei due periodi successivi il peso e l'importanza del mercato africano (ristretto per lo più al Delta del Nilo, cioè all'importante emporio di Naukratis) e di quello italico-italiota (Etruria e Magna Grecia, per lo più) che appare in forte ascesa nel periodo finale qui considerato e che lo sarà ancora di più in quelli successivi (i cui dati verranno immessi successivamente), soprattutto per la forte richiesta del mercato etrusco. Si noti che è possibile anche una scomposizione dei dati per regioni e per singole località (in questa sede si è preferito riportare il dato di interesse più generale).

Ancora, a titolo di esempio, si fornisce un raffronto quantitativo partendo questa volta dalle forme. Si sono scelte due forme piuttosto diffuse, la *kylix* (o coppa) e l'anfora (qui considerate come forme generiche: ma sono possibili anche richieste relative a forme più specifiche, come anfora a collo, ad es.). Come si vede (v. fig. 7), la *kylix* appare una forma "commerciale" per

eccellenza: in entrambi i periodi considerati più della metà di tali forme finiscono fuori della Grecia, con una forte preferenza dell'Italia per tale forma, che si accentua man mano che si procede verso la metà del secolo. Da rilevare anche l'improvviso *exploit* commerciale delle anfore nella seconda metà del VI secolo (il dato è dovuto in particolare all'Etruria).

Si dà ancora una dato, relativo questa volta ad una regione, l'Etruria, in rapporto alle forme (v. fig. 8). Appare evidente la netta propensione degli etruschi all'acquisto di anfore, in contrasto, ad es., con le preferenze della Magna Grecia, rivolte di preferenza verso le coppe.

Infine, per concludere, un dato relativo ai soggetti. Si è scelto uno dei personaggi mitologici più popolari nella ceramica attica di questo periodo: Eracle. La tavola della sua distribuzione per regioni (che qui omettiamo per brevità) evidenzia al primo posto l'Etruria, con quasi un terzo delle occorrenze di tale personaggio (32,26%), seguita dall'Attica (11,29%). Il dato appare interessante, anche se l'apparente propensione degli Etruschi verso tale soggetto va piuttosto interpretata come propensione verso i vasi dotati di scene figurate in genere (specie le "anfore tirreniche"), ove più frequente appare l'iconografia di Eracle.

Oltre allo specifico campo di indagini relative al commercio della ceramica, questi dati possono essere utilizzati anche come preziosi indici per la storia economico-sociale in genere della regione. A titolo di esempio basterà ricordare alcune significative concordanze fra i dati noti in base alle fonti letterarie (peraltro lacunose) e quanto emerge dai dati quantitativi relativi alla ceramica attica. I provvedimenti soloniani, ad es., come ci attestano le fonti letterarie, mirarono ad un forte rafforzamento del potere centrale della Polis, contro le tendenze centrifughe dell'aristocrazia sparsa nel territorio. La città venne dotata di un'agorà e venne rivitalizzata la funzione dell'acropoli. L'arcontato di Solone, lo ricordiamo, si situa al 594 a.C.. In precedenza, nel corso del VII secolo, l'Attica soffriva di un forte arretramento culturale rispetto ai centri più evoluti del mondo greco (come Corinto) o dell'Asia Minore, con una predominanza dell'aristocrazia terriera avida e provinciale. Considerando i dati raccolti relativi alla ceramica attica come indici quantitativi di storia sociale, si vede confermato dal dato numerico la forte importanza assunta dal centro cittadino (Atene) rispetto al suo territorio, che perde sempre più terreno come fonte di rinvenimento della ceramica attica rispetto alla città. Solo officine dalla produzione sempre più scadente continuano a lavorare nei villaggi (es. a Vari) al servizio dell'aristocrazia locale. Per tutte le altre officine appare ormai evidente, al contrario, il richiamo del centro cittadino.

Altro dato interessante è l'improvviso emergere della coppa *(kylix)* in età soloniana e la sua netta prevalenza nella città rispetto al territorio, a riprova del formarsi di un'aristocrazia cittadina che ha ormai recepito le usanze raffinate del simposio e ne adotta, da Corinto, l'iconografia. Si ricorda che in precedenza, nel VII secolo, persisteva ancora l'arcaica usanza, da parte dell'aristocrazia terriera, di ordinare grandi vasi funerari da adibire come segnacolo.

I dati quantitativi attestano d'altra parte ancora una certa esilità del disegno politico di una Polis capace di imporsi sulle autonomie locali (non a caso sarà necessaria successivamente la tirannide di Pisistrato per rafforzare un disegno politico unitario). In effetti, se si esaminano dal punto di vista quantitativo i soggetti sui vasi attici di tale periodo, si nota una evidente dispersione. Mancano ancora quei soggetti co-

Fig. 8 - Le forme della ceramica attica esportate in Etruria (fino alla metà del VI sec. a.C.)

LE FORME ESPORTATE IN ETRURIA (dal 630 al 550 a.C.)

ANFORA	70,06 %
KYLIX (COPPA)	15,29 %
HYDRIA	6,37 %
DINOS	2,55 %
OLPE	1,27 %
CRATERE	1,27 %
COPERCHI	1,27 %
LEKYTHOS	1,27 %
OINOCHOE	0,65 %

me Teseo o Athena, funzionali al disegno di una Polis unitaria raccolta intorno ai suoi eroi nazionali, mentre appare evidente un'assimilazione abbastanza indiscriminata da altre fonti iconografiche (Corinto, soprattutto).

5. Conclusioni.

Riassumendo e concludendo. Le società della Grecia e di Roma antiche, non diversamente da tutte le altre società storiche, pur nella riconosciuta eccellenza qualitativa della loro produzione materiale, hanno disseminato la loro esistenza storica di elementi quantitativi relativi ai più diversi ordini e alle più diverse situazioni della loro organizzazione (sociale, economica, informativa). Reticoli teorici relativi sia all'aspetto della comunicazione sia a quello della storia economico-sociale, benché ancora suscettibili (soprattutto il primo) di integrazioni e approfondimenti, sono stati tracciati. Non tutte le caselle dei reticoli consentono al momento, è doveroso ammetterlo, delle procedure formalizzate: nondimeno è in questa direzione che sembra procedere lo sviluppo scientifico della disciplina. Rinunciare a estendere a tali società metodologie quantitative in nome di un residuo pregiudizio estetico o classicistico sarebbe, più che un anacronismo, un alibi.

APPENDICE

Le ricerche in corso

Non esiste al momento un censimento preciso delle ricerche in corso in archeologia classica nelle Università e nei centri di ricerca più importanti.

È per questo che giunge opportuno, proprio quando questo lavoro va in stampa, il convegno "Informatica e Archeologia Classica" in programma a Lecce nei giorni 12-13 maggio 1986 presso la locale Università, e ai cui Atti rimandiamo per una informazione più estesa e puntuale. Diamo intanto qui di seguito notizia delle principali attività di ricerca, tramite l'uso del computer, relative all'archeologia classica o ad ambiti archeologici ad essa prossimi.

ATENE, Scuola Archeologica Italiana
curatore: *G. Bertucci*
tema: Realizzazione di banche dati archeologiche su personal computer

BOLOGNA, Università - Istituto di Archeologia
curatori: *G. Gualandi - R. Rosati*
tema: La catalogazione elettronica della ceramica attica in Etruria Padana. Ricerca a base regionale

BOLOGNA, Università - Istituto di Archeologia
curatori: *R. Rosati - M.P. Guermandi - L. Quartili*
tema: L'esportazione della ceramica attica. [Completa fino al 550 a.C.]

BOLOGNA, Università - Istituto di Archeologia
curatori: *D. Vitali - A. Gottarelli*
tema: Archiviazione di immagini dello scavo di Monterenzio (archeologia celtica)

LECCE, Università
curatori: *G. Semeraro - C. Mangia*
tema: Programma per la gestione dei dati di scavo (l'esperienza di Otranto)

MILANO, Musei Civici
curatore: *E. Arslan*
tema: Numismatica e informatica

PARIGI, Università Parigi X
curatori: *A. Ginouvés, A.M. Guimier-Sorbets*
tema: Realizzazione di banche dati relative ad ambiti diversi di archeologia classica

PISA, Scuola Normale Superiore
curatore: *P.E. Arias* e collaboratori
tema: Trasposizione su elaboratore elettronico dei dati del Corpus Vasorum Antiquorum relativi alla ceramica attica (completa).

ROMA, Università Roma I
curatore: *P. Sommella*
tema: La cartografia archeologica computerizzata

ROMA, CNR
curatore: *P. Moscati*
tema: Applicazione di metodi matematico-statistici allo studio degli specchi etruschi

ROMA, Ufficio Centrale Catalogo
curatore: *F. Parise Badoni*
tema: Elaborazione automatica di pitture e pavimenti pompeiani

TORINO, Università
curatore: *A. Michelis*
tema: Verifica di proposte di restauro e coperture edifici di Pompei

TORINO, Università
curatore: *C. Masieri*
tema: Classificazione automatica di reperti archeologici

TORINO, Sopr. archeologica Piemonte
curatore: *G. Salva*
tema: Rilievo architettonico automatizzato e simulazioni di ricostruzione nelle indagini archeologiche

URBINO, Università
curatore: *C. Gasparri*
tema: Banca dati sulle collezioni di scultura antica.

Bibliografia

A) ARCHEOLOGIA E CALCOLATORI

[1] M. BORILLO, *Formal Procedures and the Use of Computers in Archaeology,* in *Norwegian Archaeological Review,* 4, (1971), pp. 2-27.

[2] J.C. GARDIN (ed.), *Archéologie et Calculateurs,* CNRS, Paris 1970

[3] J.C. GARDIN, *Archéologie et Calculateurs, nouvelles perspectives,* in *Revue Internationale de Sciences Sociales,* XXIII, 2 (1971), pp. 204-218.

[4] J.C. GARDIN, *Code pour l'analyse des formes de poteries,* CNRS, Paris, 1976

[5] J.E. DORAN - F.R. HUDSON, *Mathematics and Computers in Archaeology,* Edinburgh, 1976.

B) DIBATTITO METODOLOGICO E NEW ARCHAEOLOGY

[1] J. DEETZ, *Invitation to Archaeology,* New York 1965

[2] K.C. Chang, *Rethinking Archaeology,* New York 1967

[3] D.L. CLARKE, *Analytical Archaeology,* London 1968

[4] S.R. BINFORD - L.R. BINFORD, *New Perspectives in Archaeology,* Chicago 1968

[5] J.M. FRITZ - F.T. PLOG, *The Nature of Archaeological Explanation,* in *American Antiquity,* 35, 4, (1970), pagg. 405-412.

[6] D.L. CLARKE (ed.), *Models in Archaeology,* London 1972.

[7] C. RENFREW (ed.), *The Explanation of Culture Change. Models in Prehistory,* London 1973.

C) ARCHEOLOGIA E METODI QUANTITATIVI

[1] M. BORILLO - A. GUENOCHE - W.F. DE LA VEGA, *Raisonnement et méthodes mathématiques in archéologie, CNRS,* Paris 1977

[2] M. BORILLO, *Archéologie et Calcul,* Paris 1978

[3] C. ORTON, *Mathematics in Archaelogy,* Cambridge 1980

D) SIMULAZIONI

[1] I. HODDER (ed.), *Simulations Studies in Archaelogy,* Cambridge 1978

E) BANCHE DATI

[1] M. BORILLO - J.C. GARDIN (edd.), *Les Banques de données archéologiques,* CNRS, Paris 1974.

[2] S.W. GAINS (ed.), *Data Bank Applications in Archaeology,* Tucson, 1981.

F) SPATIAL ARCHAEOLOGY

[1] I. HODDER, *The Spatial Organisation of Culture,* London 1978.

Spedizione in abbonamento postale a tariffa intera/Milano